**Tab. 1: Human endogenous retroviral elements**

| | HERV family | Copy number | % of genome |
|---|---|---|---|
| Class I HERVs (type C- related HERVs) | **HERV-ERI** | | 0.07 % |
| | HERV-E (4-1, ERVA, NP-2) | 35 - 50 | |
| | HERV-E LTR | 500 - 600 | |
| | 51-1 | 35 - 50 | |
| | ERV1 | 10 - 15 | |
| | HERV-R (ERV3) | 10 | |
| | RRHERV-I | 20 | |
| | **HERV-T** (S71, CRTK1, CRTK6) | 50 - 60 | |
| | HERV-T LTR | 150 - 200 | |
| | **ERV-FRD** | 8 | |
| | **HERV-RW** | | 0.2% |
| | HERV-W (MSRV) | 25 - 50 | |
| | HERV-R (ERV9) | 30 - 40 | |
| | ERV9 LTR | 3000 - 4000 | |
| | **HERV-P** (HuERS-P, HuRRS-P) | 50 - 90 | 0.01% |
| | **HERV-IP** | | 0.01% |
| | HERV-I (RTVL-I) | 25 - 50 | |
| | HERV-IP-T47D (ERV-FTD) | 35 | |
| | HERV-IP LTR | 1800 - 2000 | |
| | **HERV-HF** | | 0.2% |
| | HERV-H (RTVL-H, RGH) | 900 - 1000 | |
| | HERV-F | 16 | |
| | HERV-H-LTR | 1000 | |
| Class II HERVs (type A-, B-, and D-related HERVs) | **HERV-K** | | 0.5% |
| | *HERV-K(HML-1)* | 10 - 20 | |
| | *HERV-K(HML-2)* | 30 - 50 | |
| | HERV-K10 | | |
| | HERV-K-HTDV | | |
| | HERV-K-IDDM | | |
| | *HERV-K(HML3)* | 25 | |
| | *HERV-K(HML-4)* | 6 | |
| | HERV-K-T47D | | |
| | *HERV-K(HML-6)* | 30 - 40 | |
| | HERV-K-HML-6p | | |
| | HERV-K-HML-6.17 | | |
| | *HERV-K(HML-7)* | ? | |
| | HERV-K-NMW7 | | |
| | *HERV-K(HML-8)* | ? | |
| | *HERV-K(HML-9)* | ? | |
| | HERV-K-NMW9 | | |
| | *HERV-K(HML-10)* | 10 - 50 | |
| | HERV-KC4 | | |
| | HERV-K LTR | 10 000 - 25 000 | |
| Foamy virus-related HERVs | **HERV-L** | 100 - 200 | 0.02% |

**Tab. 2:** Primers used for the amplification of different HERV-LTR-regions

| Nr. | Primer | Sequence |
|---|---|---|
| 34 | HERV-K | ATGGCGGTTTTGTCGAA |
| 35 | HERV-K | GTTCCMTYAGTATTTATTGATC |
| 36 | HERV-K | ATGGAGCATACAATCGGG |
| 3 | HERV-K | AAGAAAAGGGGGAAATGTGGG |
| 11 | HERVKC4 | AAAGGGAGGGGGGCATG |
| 12 | HERV-KT47-D | TAAAAAGGGGGGAGATG |
| 1 | HERV-H | ATGTGAGCAACATGGCTGTTTATTTC |
| 2 | HERV-H | TGTCAGGCCTCTGAGCCCAA |
| 39 | HERV-H | GCCATCTCGAGTGTCAGSCCTCTGAGYCYARGC |
| 37 | HERV-H | TATCTTGAATTCGKGTGAGCAAYAARRCTTTA |
| 31 | polydT | TTTTTTTTTTTTTTTTT |
| 17 | HERV-E | AAAGGGGGGGAAATATG |
| 18 | HERV-L | AGGGGTGGGACTTGCGATG |
| 19 | HERV-W | TGTTGAGATGGGGGACTGAG |
| 20 | HERV-W | GCAGTTGCAAGATTTAATAGAG |

**Tab. 3:** Analyzed HERV-LT

**A:** HERV-LTRs from different cell lines and tissues

| | Primer | Herkunft | Homology |
|---|---|---|---|
| HERV-K2 | 34/36 | T47-D | HERV-K10, M12854, 97,9 % |
| HERV-K3 | 34/36 | T47-D | HERV-K10, M12854, 98,4 % |
| HERV-K22-K32-K27-K45 | 34/36 | brain | HERV-K10, M12854, 98,6 % * |
| HERV-K30 | 3/31 | heart | HERV-K10, M12854, 97,6 % |
| HERV-K-T47D-L5 | | T47-D | MRSV, AF127229 |
| HERV-K-T47D-L50 | | T47-D | MRSV, AF127229 |
| HERV-K-T47D-L8 | | T47-D | MRSV, AF127229 |
| HERV-K-T47D-L9 | | T47-D | MRSV, AF127229 |
| HERV-K-T47D-L48 | | T47-D | MRSV, AF127229 |
| HERV-K-T47D-L20 | | T47-D | MRSV, AF127229 |
| HERV-IP-T47D | | T47-D | MRSV, AF127229 |
| HERV-T47D-W2 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-T47D-W4 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-T47D-W5 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W1 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W10 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W11 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W18 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W2 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W22 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W23 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W4 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W5 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W6 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-W8 | 19/20 | T47-D | MRSV, AF127229 |
| HERV-H1 | 1/2 | H9 | Cercopithecus aethiops ERV-H; U96012, 87,1% |
| HERV-H8 | 1/2 | HUT | HERV-H LTR18106, 84,8% |
| HERV-H13 | 1/2 | HUT | HERV-H LTR18106, 91,8% |
| HERV-H19 | 1/2 | liver | Callithrix jacchus ERV-H, 5´LTR; U96052, 92,1% |
| HERV-H31 | 1/2 | liver | HERV-H(H6) x12717, 99,8% |
| HERV-H3 | 1/31 | 85HG66 | HERV-H(H6) x12717, 100 % |
| HERV-H CL1 | 1/2 | Chang Liver | HERV-H(H6) x12717, 100 % |
| HERV-H CL2 | 1/2 | Chang Liver | HERV-H LTR18106, 84 % |
| HERV-H CL3 | 1/2 | Chang Liver | Callithrix jacchus ERV-H, 5´LTR Silva 5, U96057, 84,2 % |
| HERV-H CL4 | 1/2 | Chang Liver | HERV-H(H6) x12717, 100 % |
| HERV-H PA7 | 1/2 | Panc1 | Callithrix jacchus ERV-H, 5´LTR Silva 4, U96062, 85,7 % |
| HERV-H PA8 | 1/2 | Panc1 | Cercopithecus aethiops ERV-H, Vero 22, U96012, 87,1% |
| HERV-H PA9 | 1/2 | Panc1 | HERV-H LTR18106, 85 % |
| HERV-H PA10 | 1/2 | Panc1 | Callithrix jacchus ERV-H, 5´LTR Silva 4, U96062, 85,6 % |
| HERV-H MC14 | 1/2 | MCF7 | Cercopithecus aethiops ERV-H, Vero 22, U96012, 86,6% |
| HERV-H MC15 | 1/2 | MCF7 | Cercopithecus aethiops ERV-H, U96012, 86,6 % |
| HERV-H MC16 | 1/2 | MCF7 | Callithrix jacchus ERV-H, 5´LTR Silva 4, U96062, 87,4 % |
| HERV-H MC17 | 1/2 | MCF7 | Cercopithecus aethiops ERV-H, Vero 22, U96012, 86,6% |
| HERV-H MP20 | 1/2 | MiaPaca | Human beta globin retrovirus-like repetitive element, k01891, 92,8 % |
| HERV-H MP21 | 1/2 | MiaPaca | HERV-H LTR18106, 89,2 % |
| HERV-H MP23 | 1/2 | MiaPaca | HERV-H(H6) x12717, 99,5 % |

**B:** HERV-LTRs published in the literature

| | (bp) | reference |
|---|---|---|
| **HERV-K-pl167** | 970 | Leib-Mösch *et al*, 1993 |
| **HERV-K-T47-D** | 1200 | Seifarth *et al.*, 1998 |
| **HERV-H-H6** | 393 | Feuchter und Mager, 1990 |
| **HERV-T-S71A** | 625 | Murr, Dissertation, 1998 |
| **HERV-E** | 391 | Steele *et al.*, 1984 |
| **HERV-L** | 462 | Cordonnier *et al.*, 1995 |

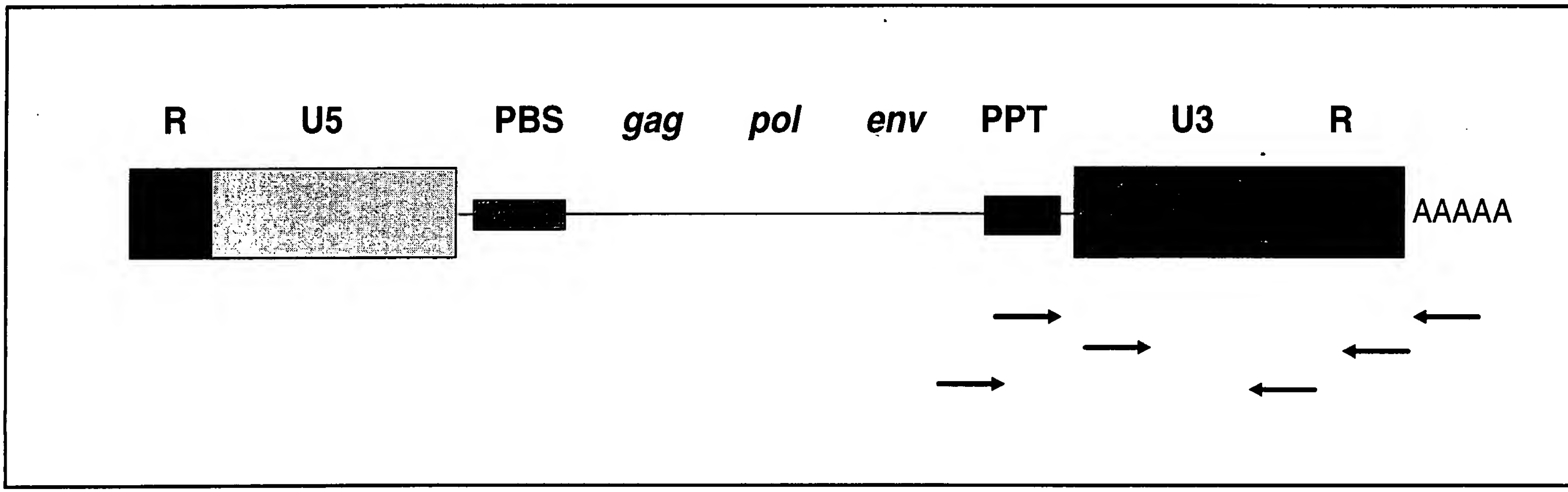



**Fig.1:** RT-PCR strategy to isolate U3/R-regions of transcribed HERVs

Abb.2a)

HERV-H-H6

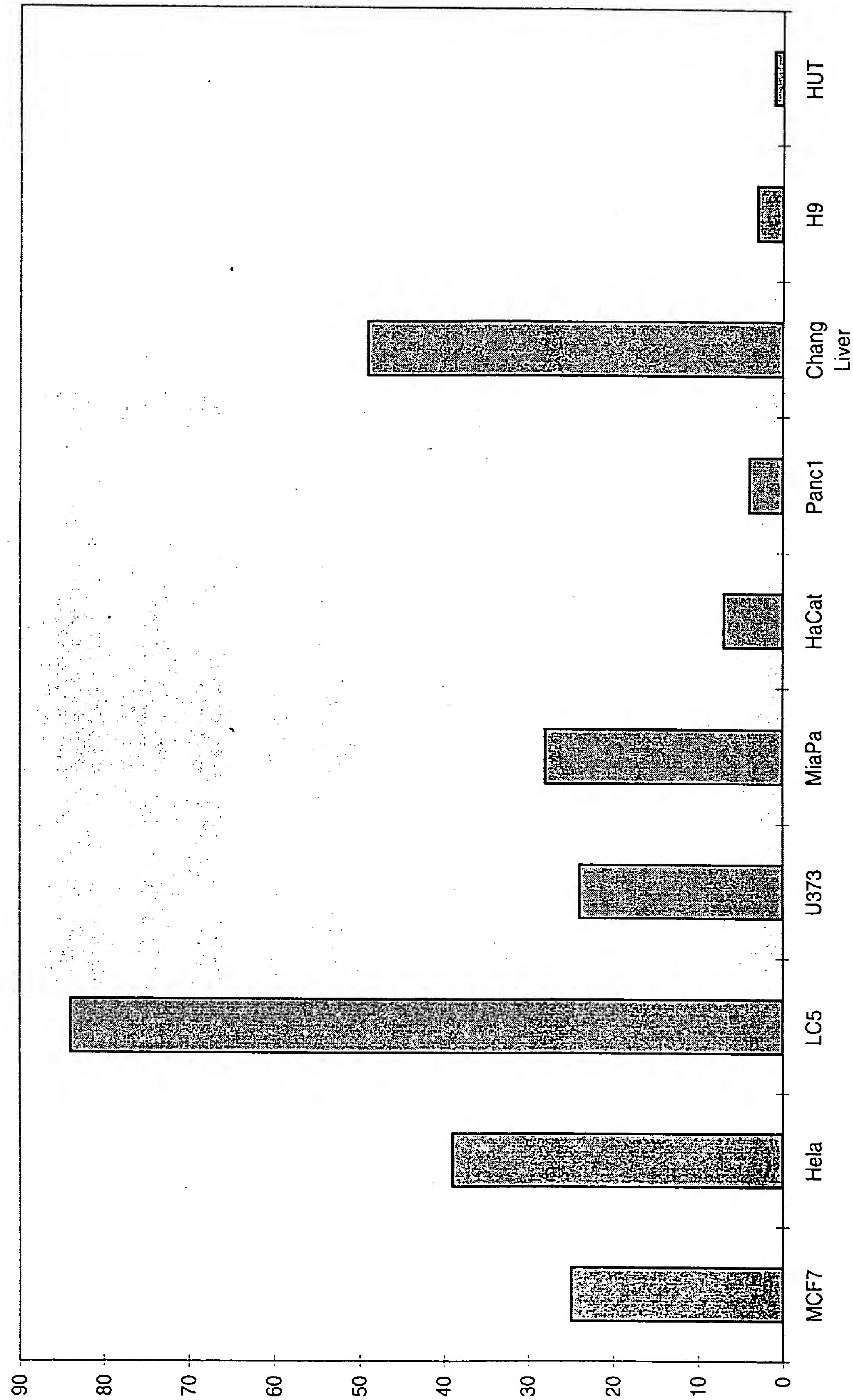


**Abb.2b)**

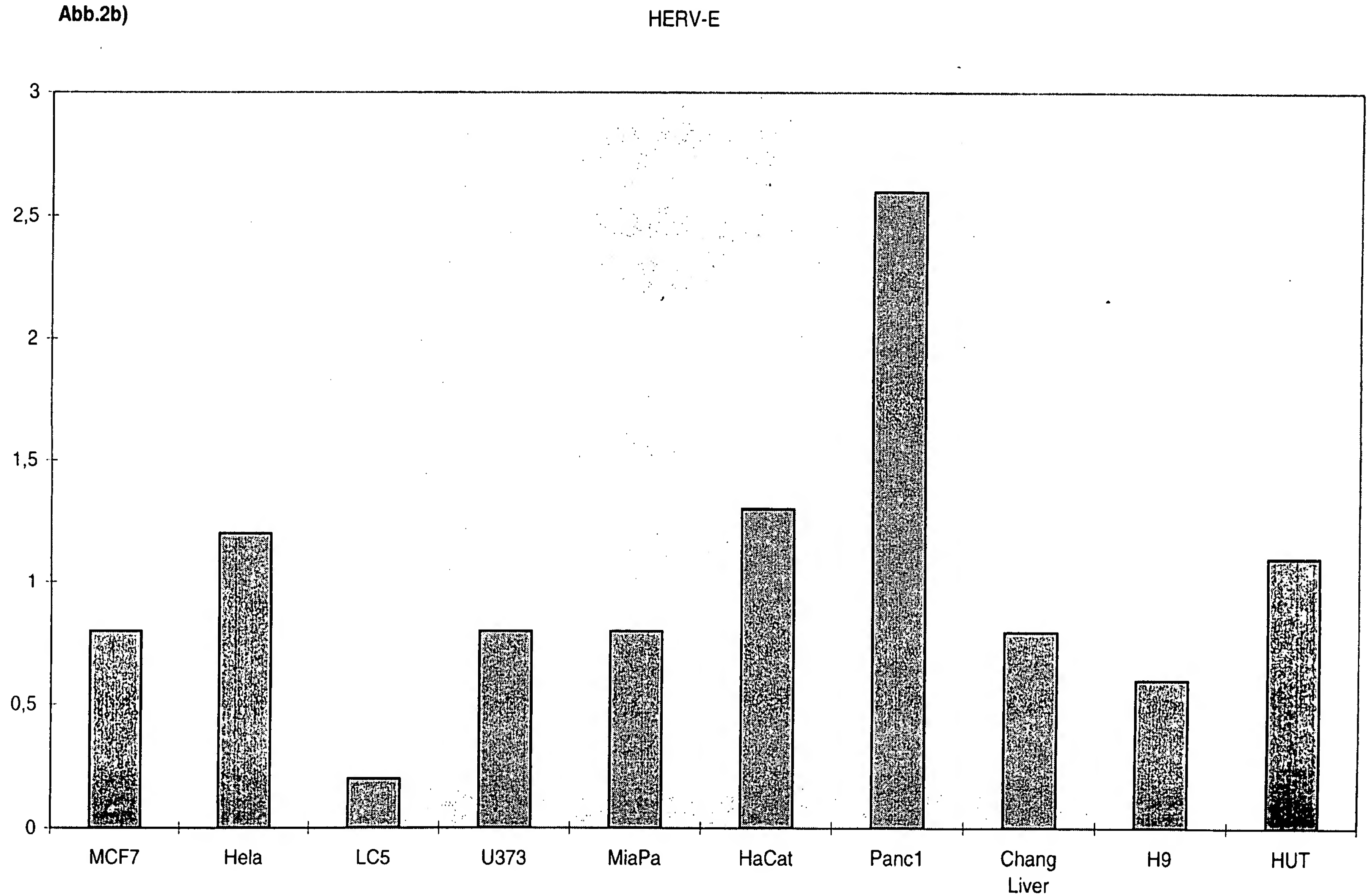

HERV-E
3
2,5
2
1,5
1
0,5
0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver
H9
HUT

Abb.2c)

HERV-Kpl167

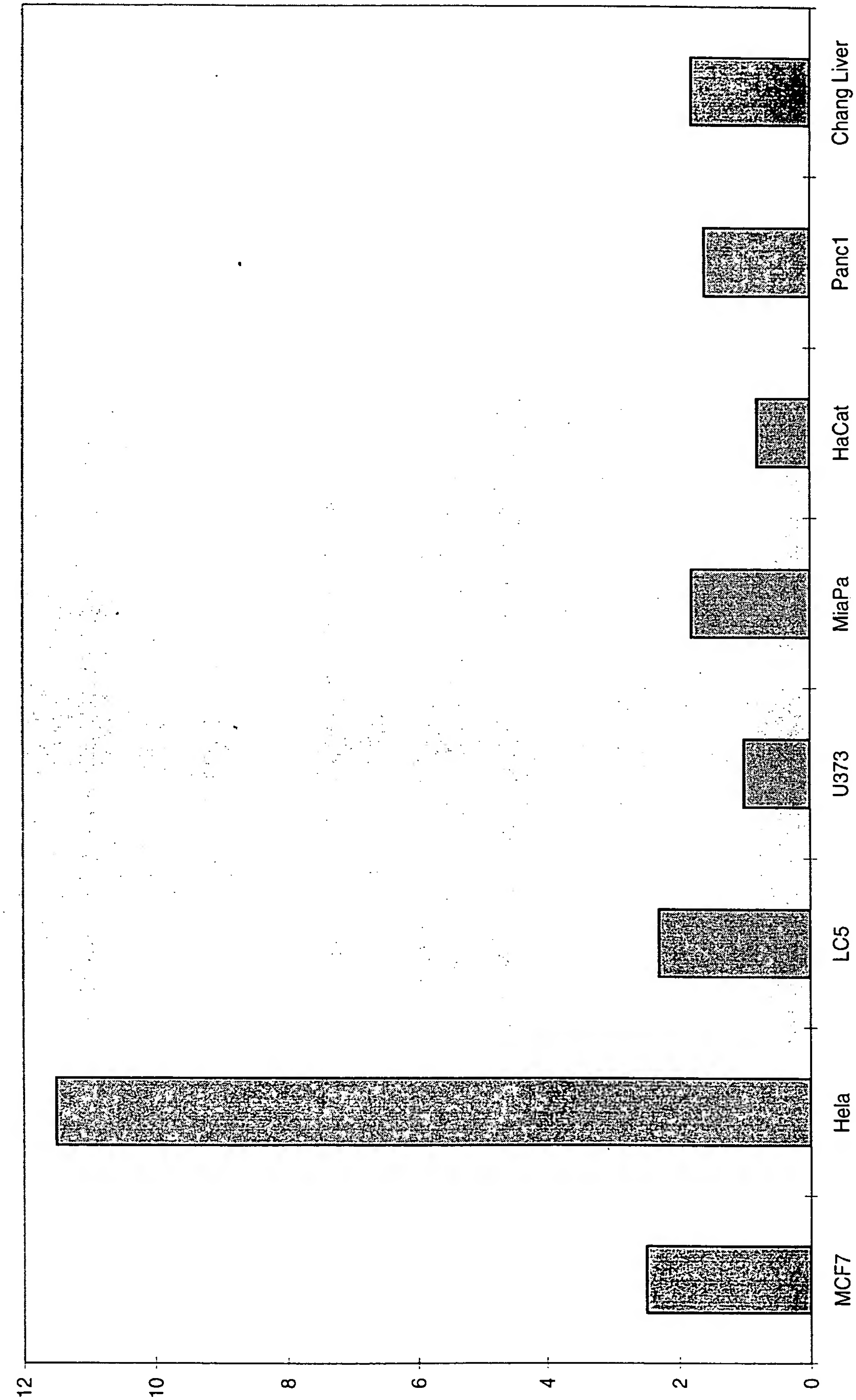


Abb.2d)

HERV-L

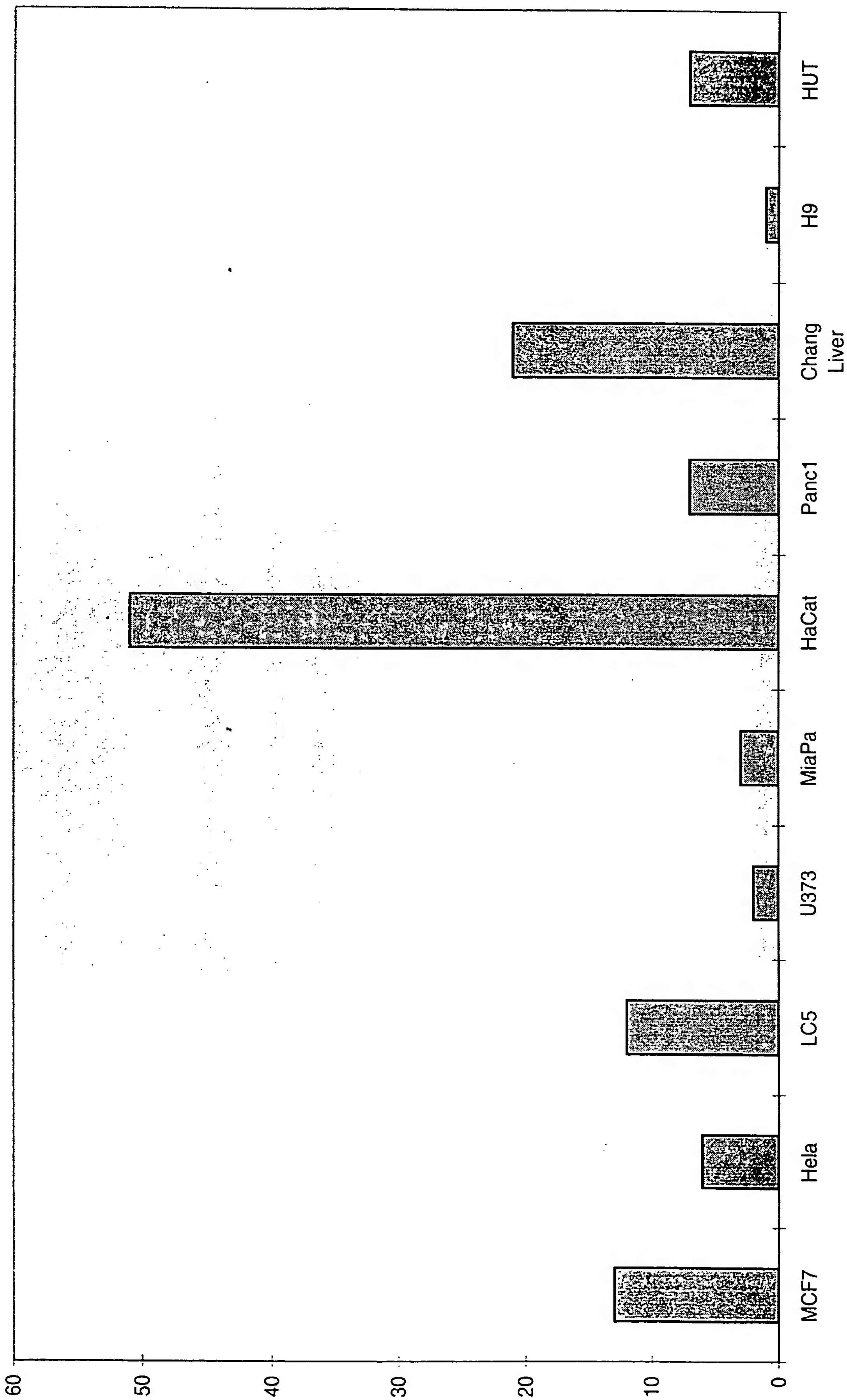

MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver
H9
HUT
0
10
20
30
40
50
60

Abb.2e)

HERV-K-T47D

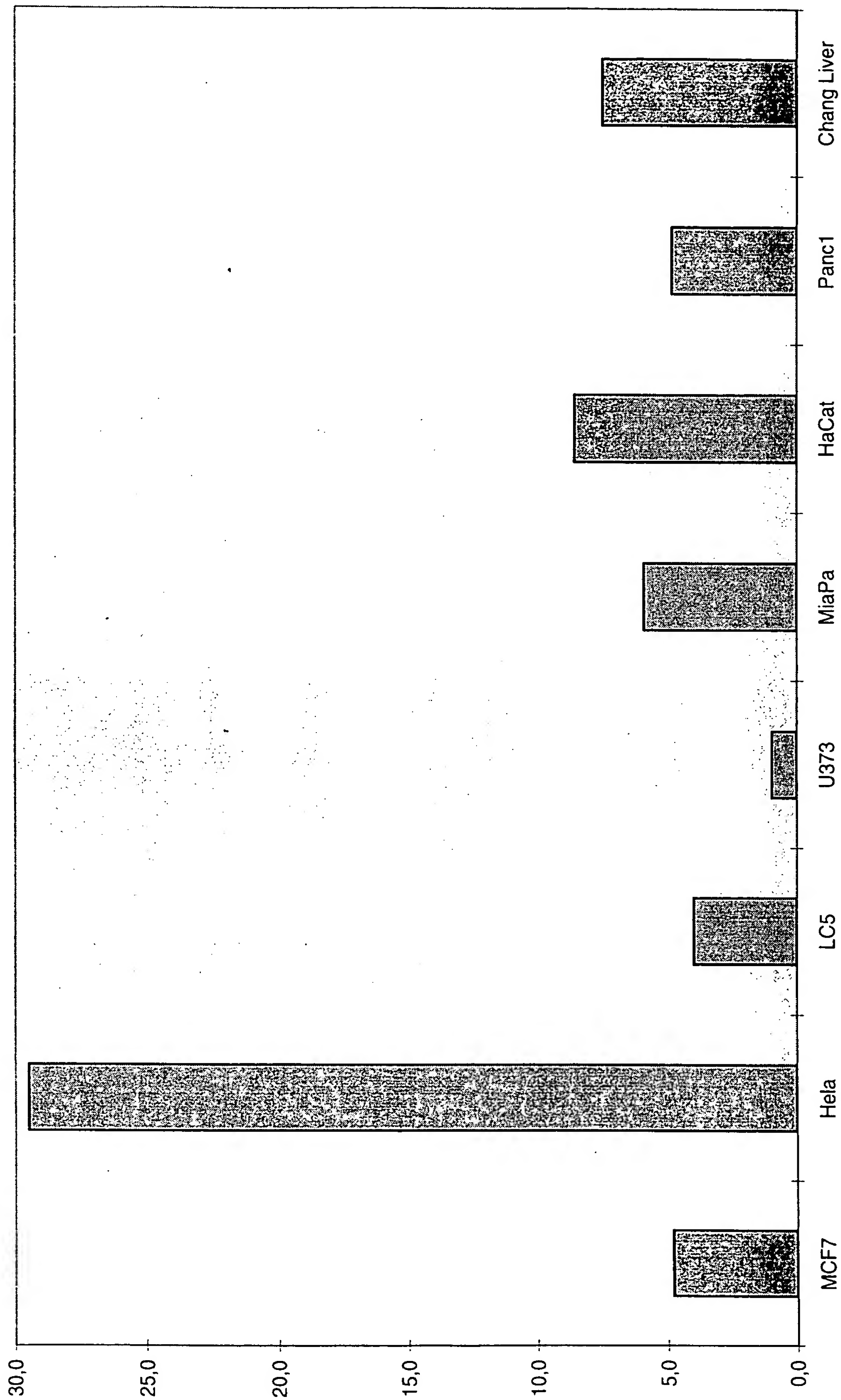
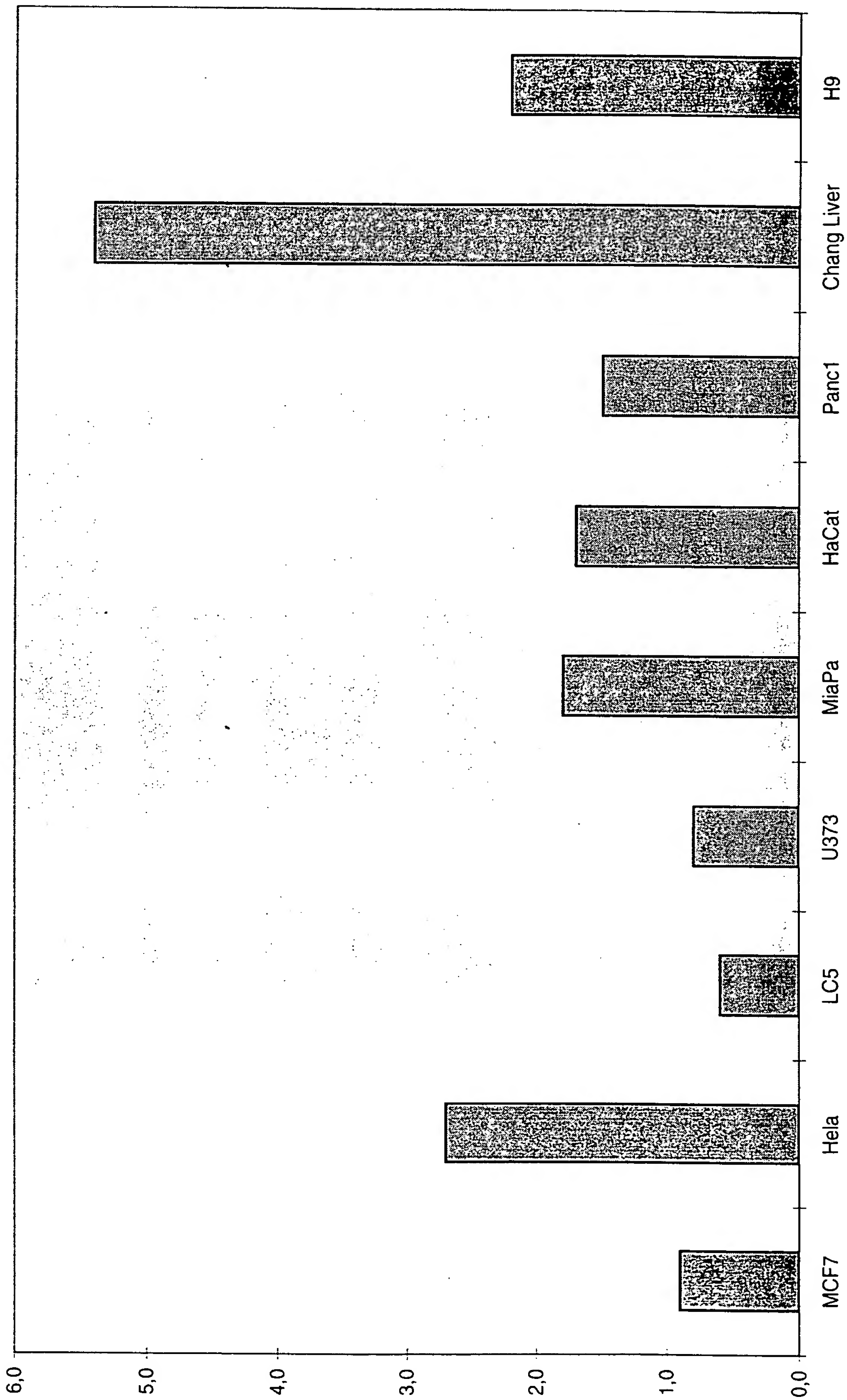


Abb.2f)

HERV-T


H9
Chang Liver
Panc1
HaCat
MiaPa
U373
LC5
Hela
MCF7
0,0
1,0
2,0
3,0
4,0
5,0
6,0

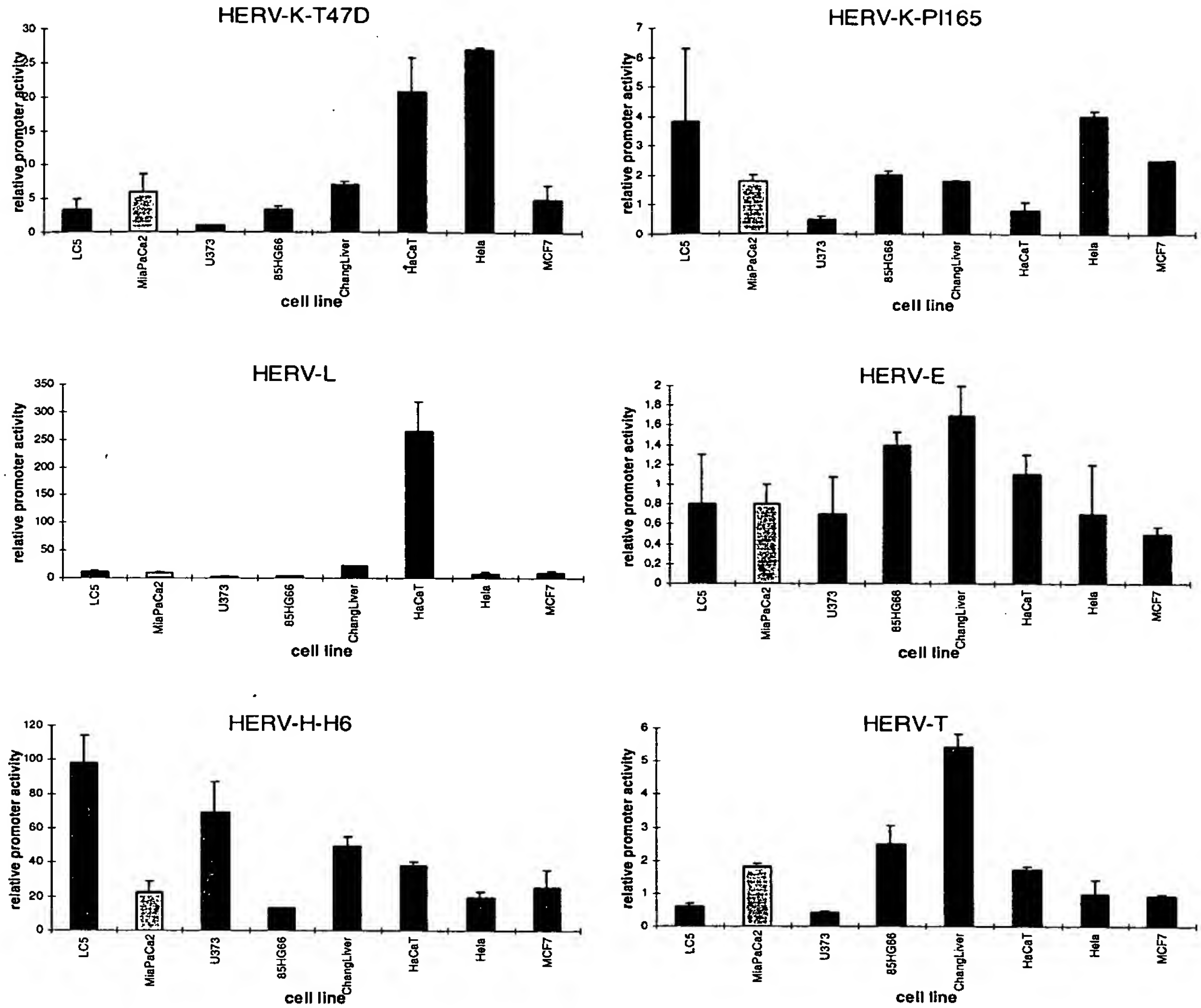
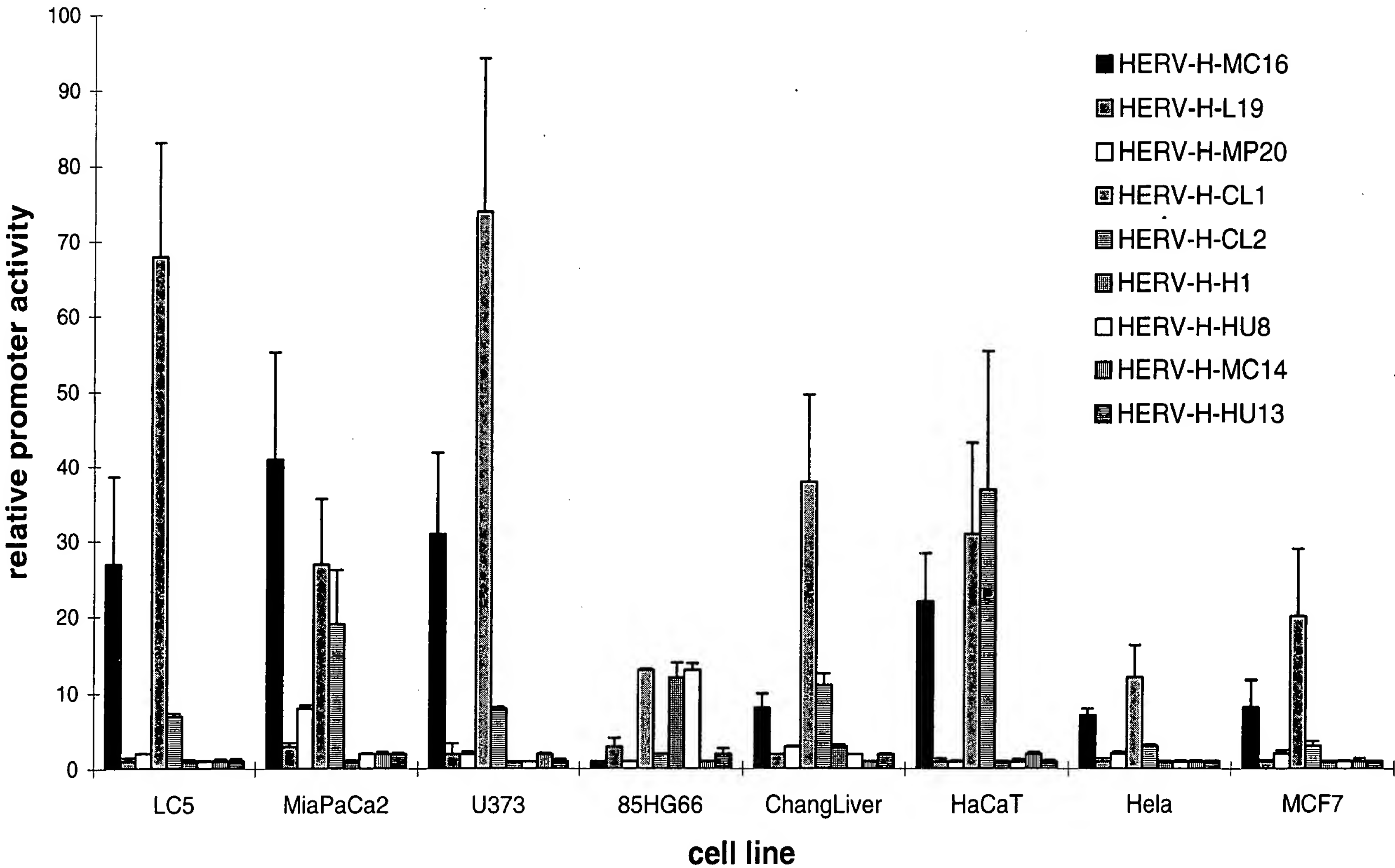


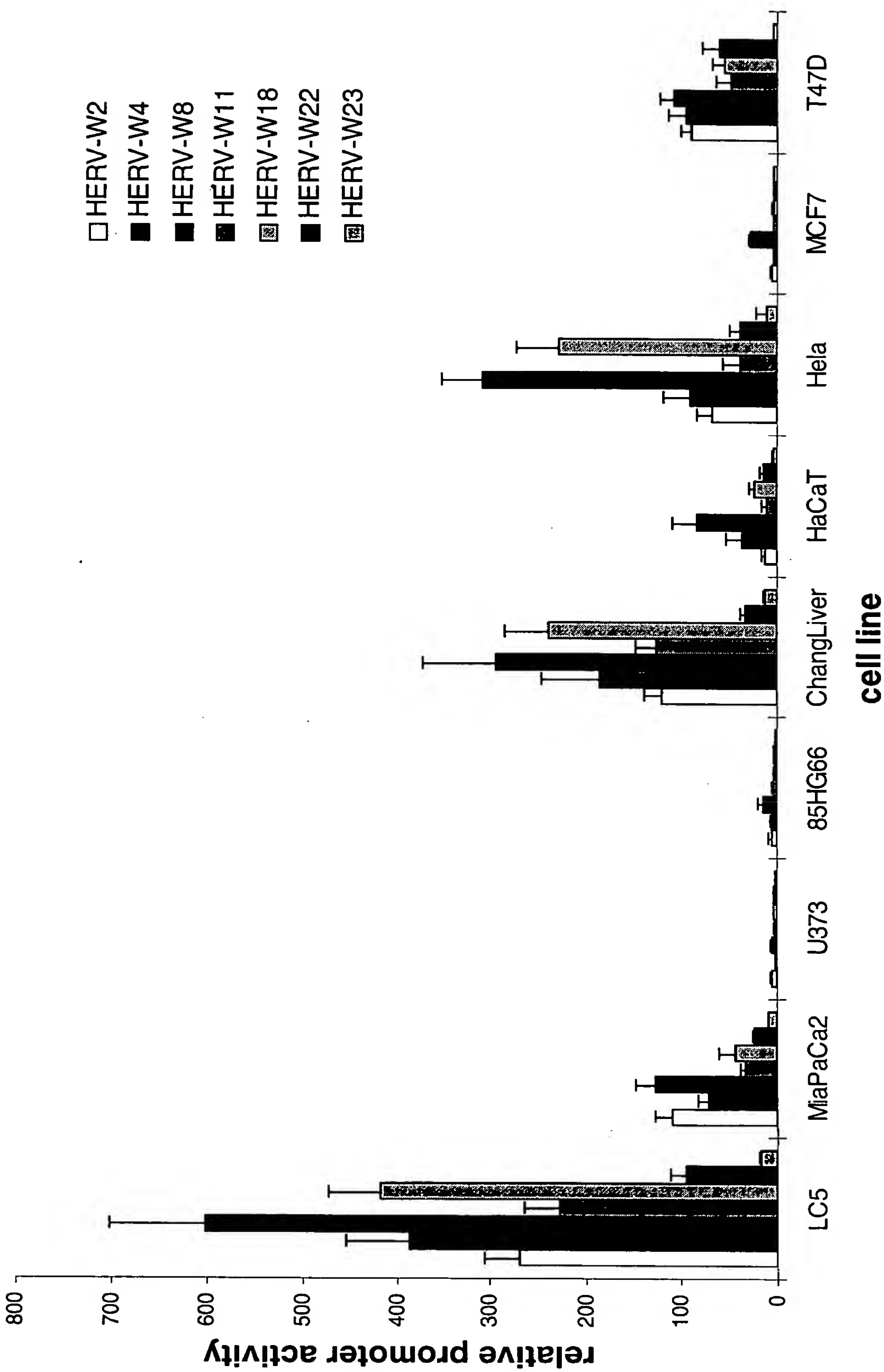

Abb. 2g: relative promoter activity of different HERV-LTRs in different cell lines

Fig. 3a



relative promoter activity
100
90
80
70
60
50
40
30
20
10
0
LC5
MiaPaCa2
U373
85HG66
ChangLiver
HaCaT
Hela
MCF7
cell line
HERV-H-MC16
HERV-H-L19
HERV-H-MP20
HERV-H-CL1
HERV-H-CL2
HERV-H-H1
HERV-H-HU8
HERV-H-MC14
HERV-H-HU13

Fig. 3b


HERV-W2
HERV-W4
HERV-W8
HERV-W11
HERV-W18
HERV-W22
HERV-W23
relative promoter activity
0
100
200
300
400
500
600
700
800
LC5
MiaPaCa2
U373
85HG66
ChangLiver
HaCaT
Hela
MCF7
T47D
cell line

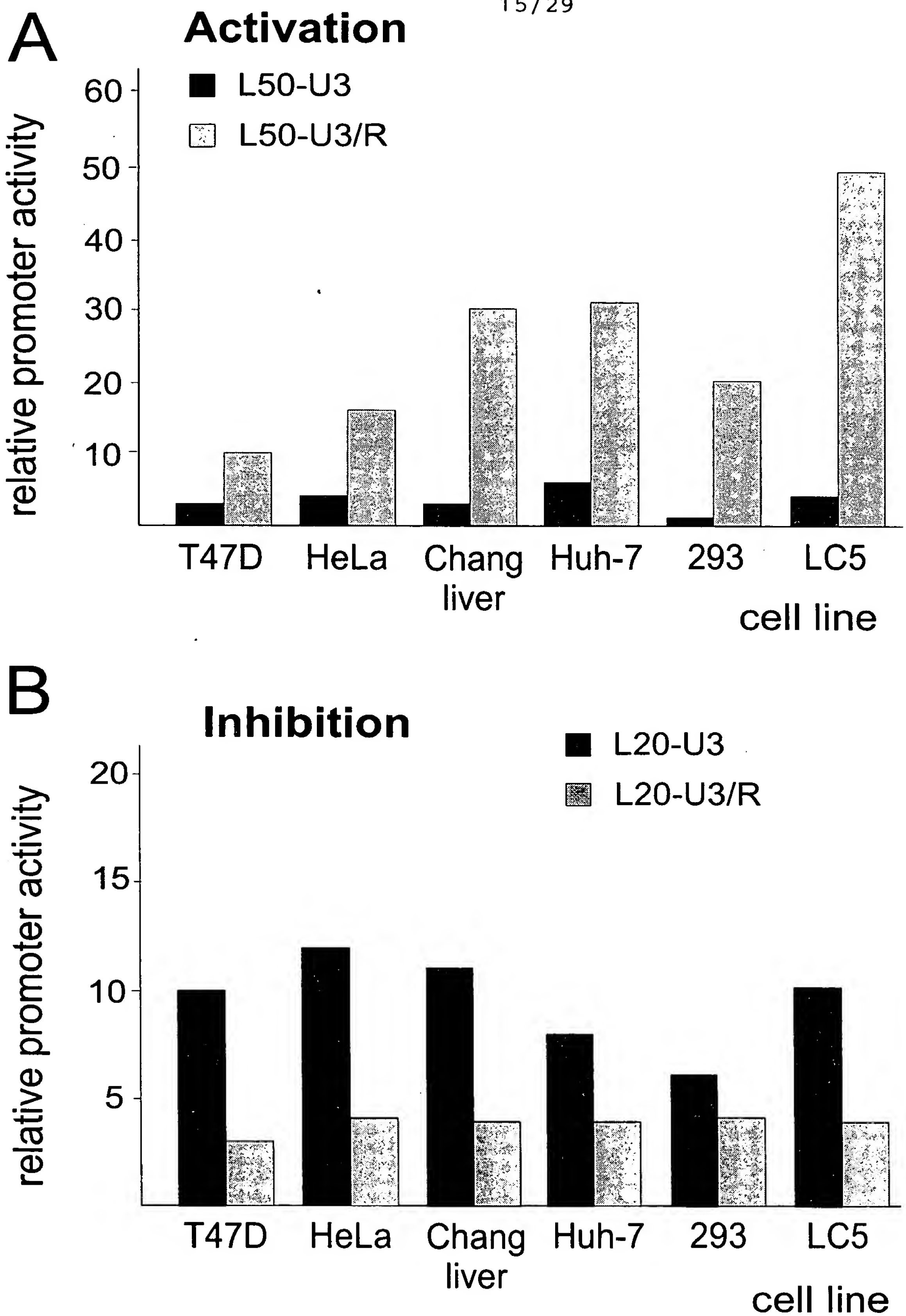



Fig. 4: LTR-R region modulates promoter activity of HERV-K-T47D related LTRs

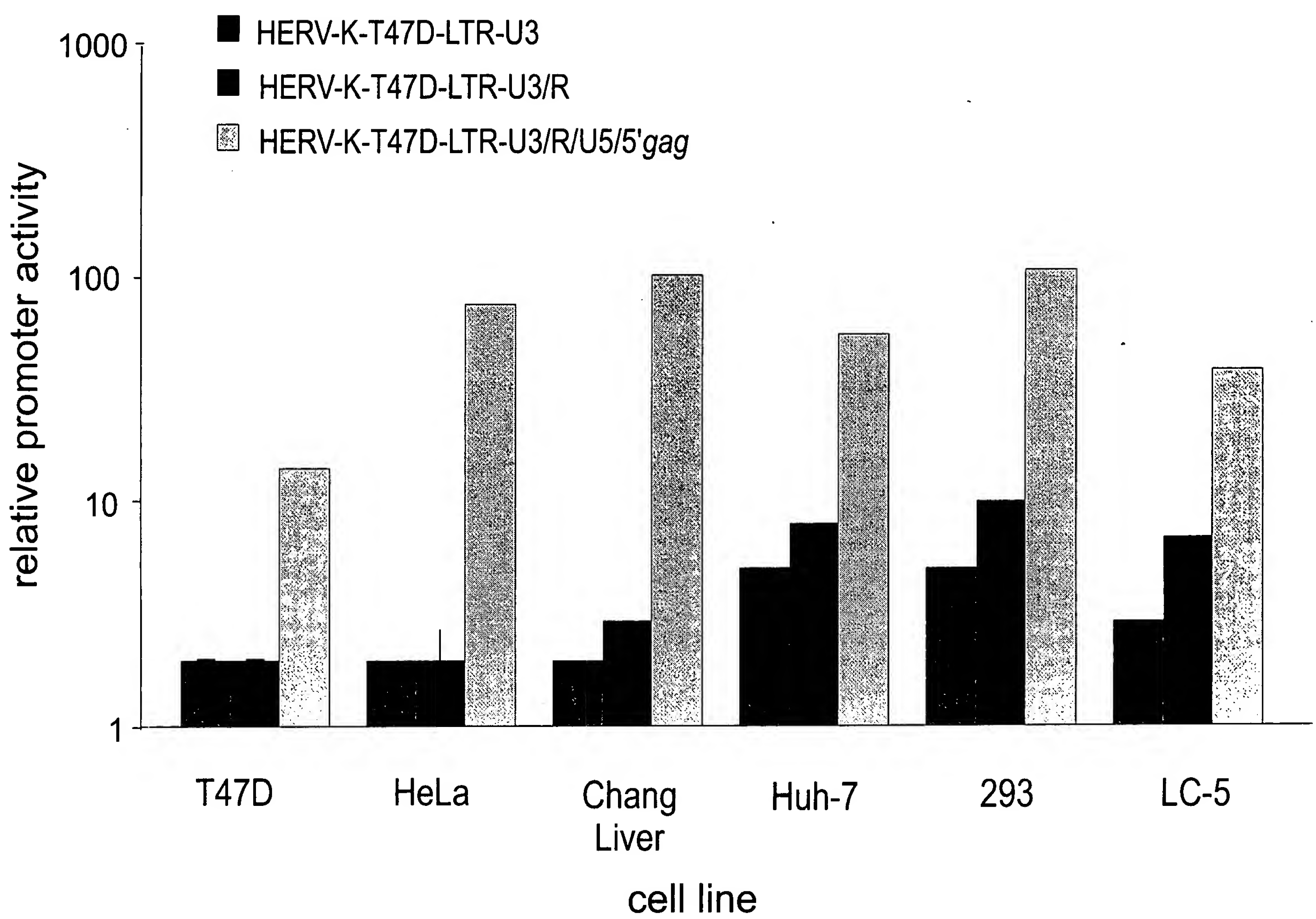



Fig. 5: Sequences downstream of LTR-R modulate promoter activity of HERV-K-T47D related LTRs

```
                           U3 <--> R
                            Inr
HERV-K-T47D  AAACCCCTCCCTGTGGTGCTGTGCTTCAATGGCCATGCTTCTTGTCCACTTTCATGTTCCTCCTGTACTCCTGGTTCCTCTTTGAAGTTCGTAGAAGATAATGGTAGAAGAAATAGTGAA  584
L5           A.....T...................A...T.CCAC.C..T............C....C.CG..........C..........AG..C..T.....GCG.......A.....T...  583
L50          A.....T...................A...T.CCAC.C..T............C....C.CG..........C..........AG..C..T.....GCG.......A.....T...  583
L8           A.....C...................A...G.ACAC.C..C............T....C.CA..........C..........TT..T..T.....GTG.......G.....G...  599
L9           A.....C...................A...G.ACAC.C..C............T....C.CA..........T..........TT..T..T.....GTG.......G.....G...  599
L48          G.....C........-----.......A...T.ACGT.C..T............C....T.CG..........C..........TT..T..T.....GCG.......A.....T...  587
L20/L49      A.....C...................G...T.ACTC.C..T............T....C.CG..........C..........TT..T..T.....GCA.......A.....T...  596

                                                                                                       NF1       GR   TFE3-S
HERV-K-T47D  AGTCTT--------TGATCTTTCTTATAAGTGCATAGAAGAAAACACTGATGTATGCCTGCCTTCCCT--CTCTGCTTCAGCTACCTAAAA-GGAAAGGCCCCCTTTCCCATGATCACAT  693
L5           .G....AAAGTCTTC..........C....G..G..........G....CATA..-.T......C..--C.....TTCG.........A.G....G....GC.TATCCTG.A....CA.  700
L50          .G....AAAGTCTTC..........C....G..G..........G....CATA..-.T......C..--C.....TTCG.........A.G....G....GC.TATCCTG.A....CA.  700
L8           .A....AAAGTGTTT..........T....G..T..........G....CATA..-.T......T..--G.....TTCA.........G.G....G....CC.TGTCCAG.G....CG.  716
L9           .A....AAAGTGTTT..........T....G..T..........G....CATA..-.T......T..--G.....TTCA.........G.G....G....CC.TGTCCAG.G....CG.  716
L48          .G....AAAGTCTTT.....----.T....T..T..........G....TGCC..-.C......T.--TC.....TTCA.........G.G....G....CG.TGTCCTG.G....GG.  700
L20/L49      .G....AAAGTATTT..........T....G..T..........G....CATA..-.T......T..ATC.....GGTG.........A.G....G....CC.TGTCCCA.G..CATGT  715

                                  AP-1              SP1            poly A
HERV-K-T47D  GACTTGCCTGACCTTATCAATCACTTGGAGGACTCACCCTCCTTACCCTG--TCCCTTTGTCTTGTATGCAATAAATATCAGCACGCCCAGCCATTCGGGGCCACT-ACTGGTCTCCGCA  810
L5           .......T.C.....G..........A..A..T.C.CT......AC...G--C...C.............................TGAC.....C.G..CAGGG.....-A.T.......GCG  817
L50          .......T.C.....G..........A..A..T.C.CC......AC...G--C...C.............................TGAC.....C.G..CAGGG.....-A.T.......GCG  817
L8           .......T.C.....G..........A..A..T.C.CC......AC...G--C...C.............................TGCA.....C.T..CGG...-....TA.C.......ACG  833
L9           .......T.C.....G..........A..A..T.C.CC......AC...G--C...C.............................TGCA.....C.T..CGGGG.....TA.C.......ACG  834
L48          .......T.C.....G..........A..A..C.G.CC......AT...G--C...C.............................CGAG.....C.G..CAGGG.....-A.C.......GTG  817
L20/L49      G......T.C.....----.......A..A..T.C.TC......AC...GCGC...C.C...........................CACG.....T.G..TGAGG.....-G.C.......GCG  830

                          R <--> U5
HERV-K-T47D  ACTTGGTGGTAGTGGTACCCTGGGCCCAGCTGTTTTCTCTTTA
L5           T...GA..........CA..CC......---------------
L50          T...GA..........CA..CC.......G.G..T.T......
L8           T...GG.........-C..CC.......C.G..T.C......
```

Fig. 6: Regulatory elements in the R region of HERV-K-T47D LTRs

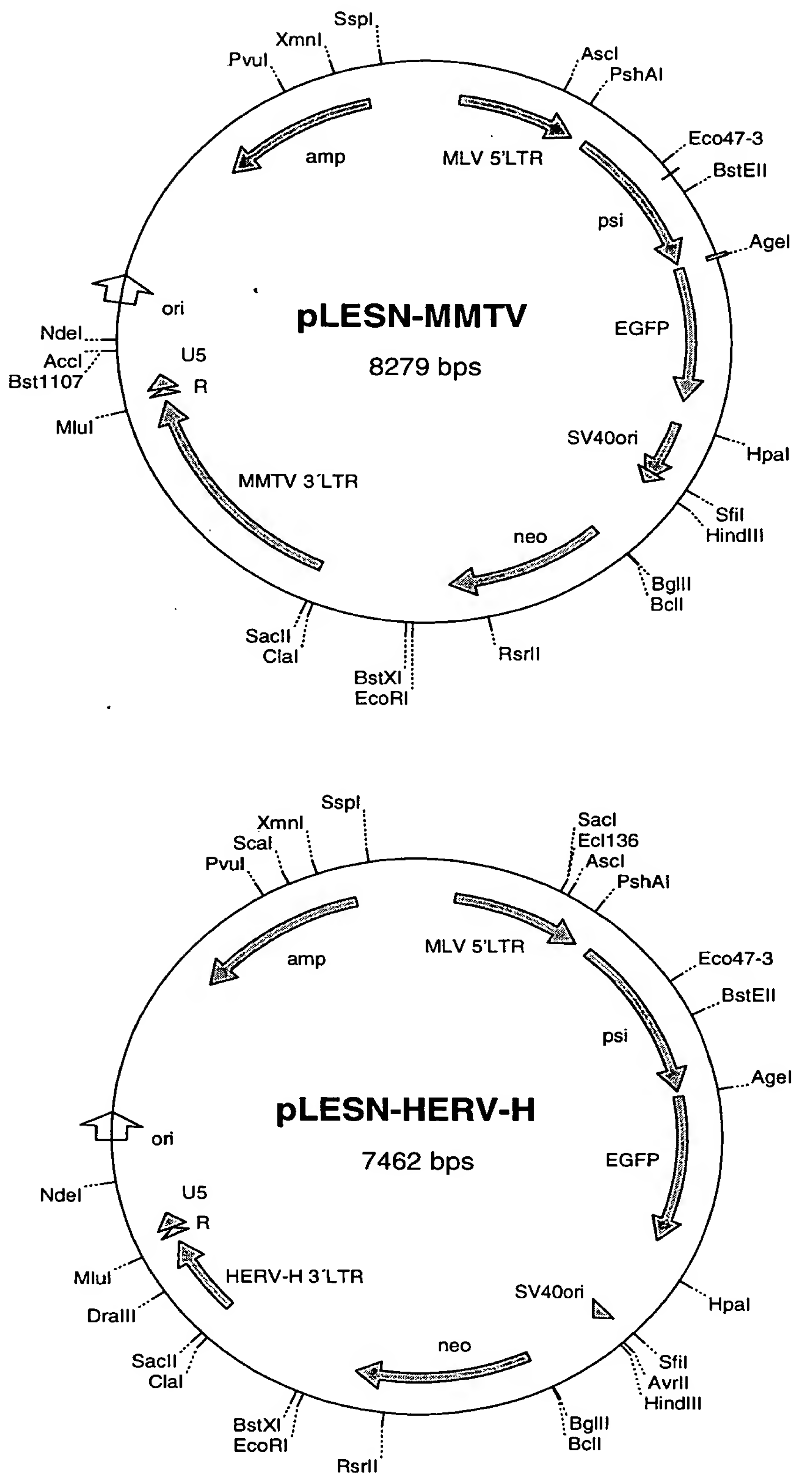
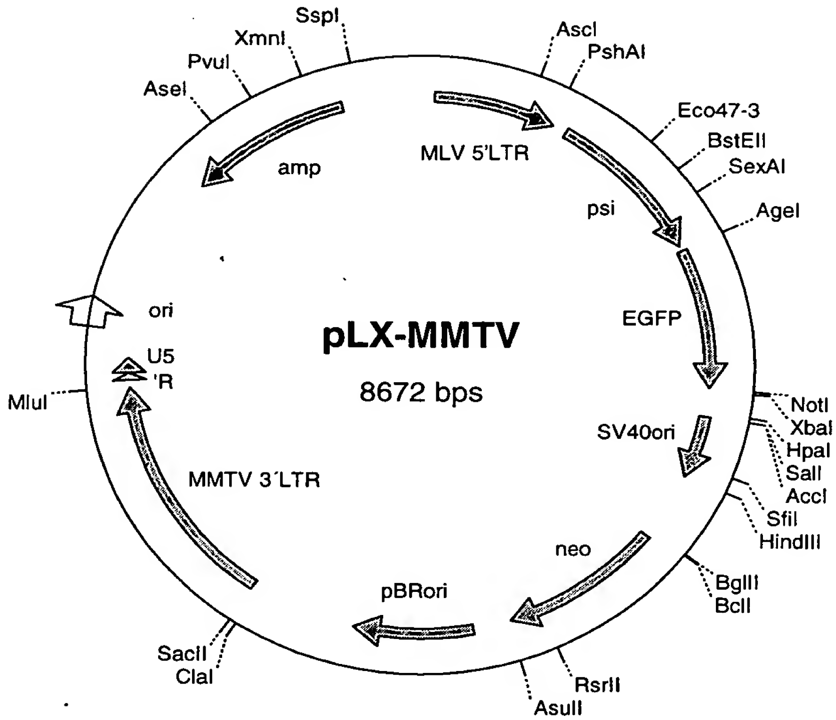



**Fig.7:** Retroviral ProCon vectors pLESN-MMTV and pLESN-HERV-H



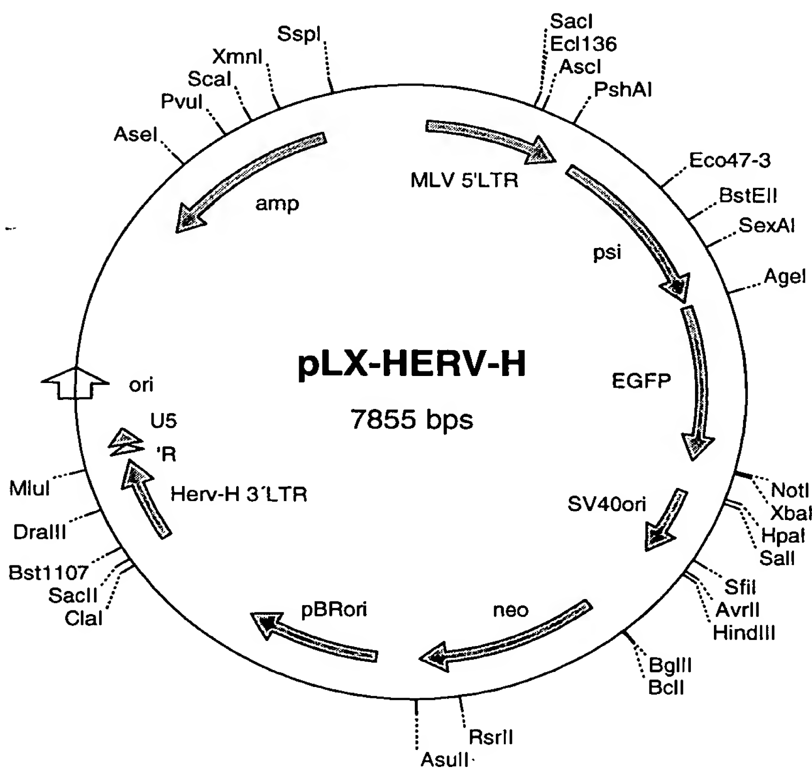
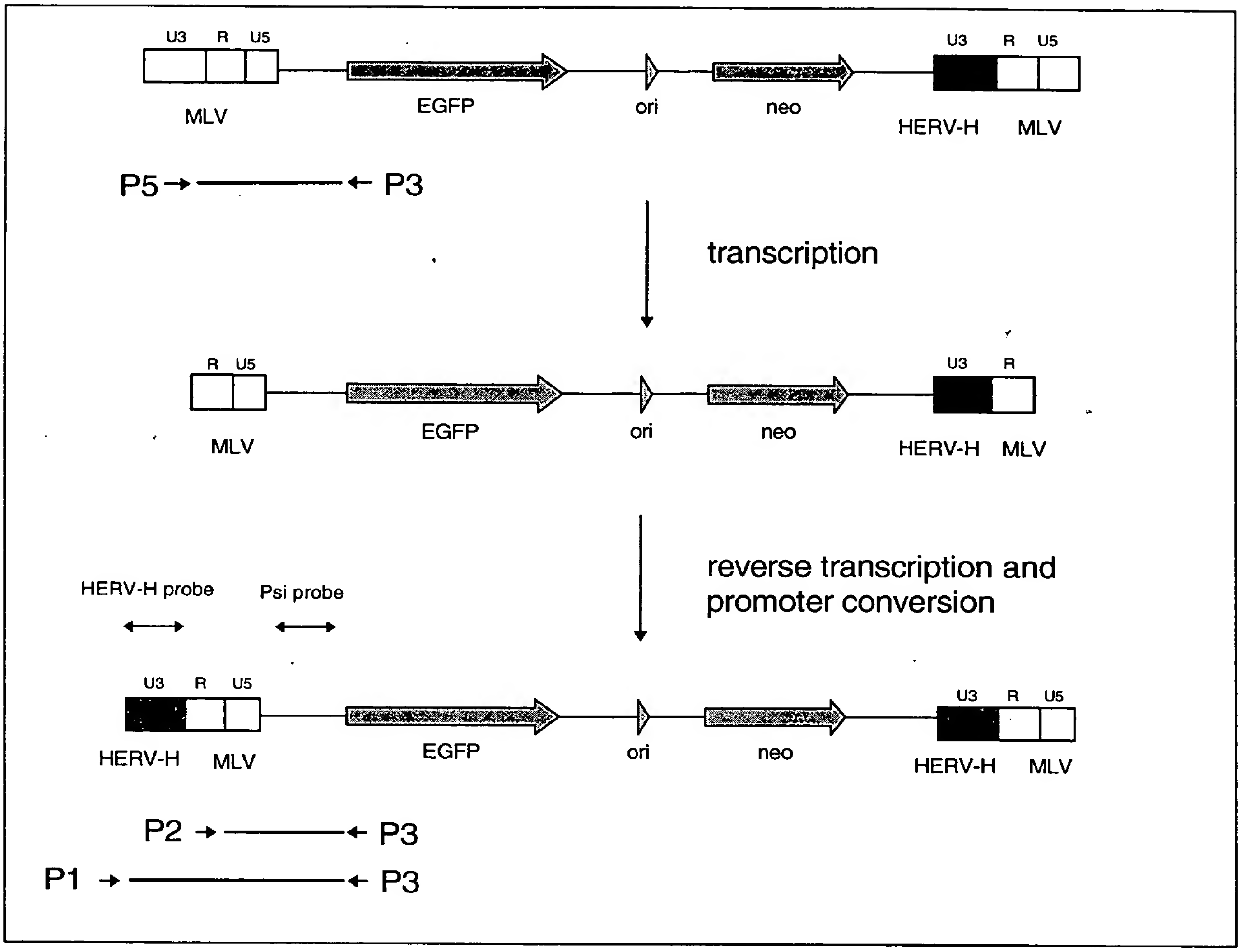




**Fig.8:** Retroviral ProCon vectors pLX-MMTV and pLX-HERV-H

a)




b)

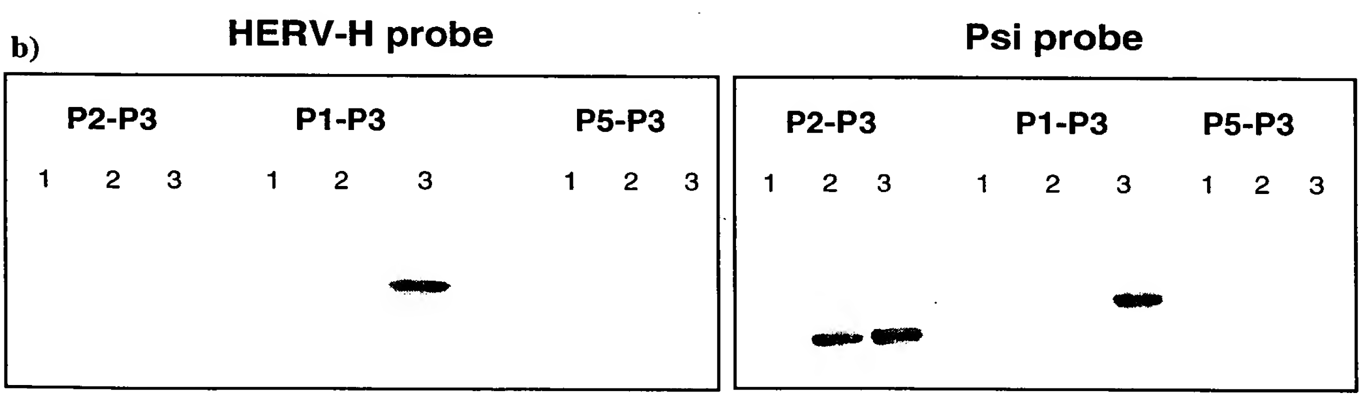
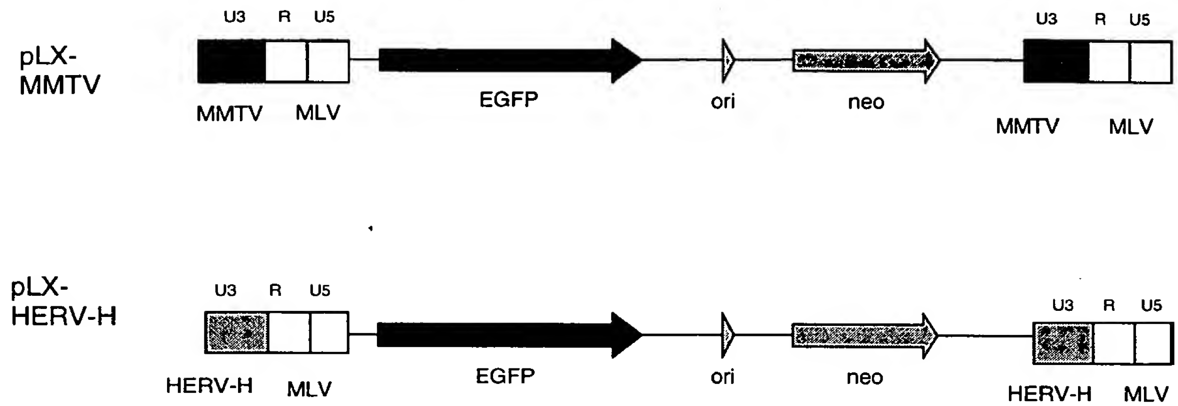



**Fig. 9:** a) Promoter conversion of the hybrid ProCon vectors
b) Demonstration of the correct promoter conversion with PCR and hybridization with a HERV-H and a psi probe (1:CK; 2:CK-pLX-MMTV; 3:CK-pLX-HERV-H)

a)




b)

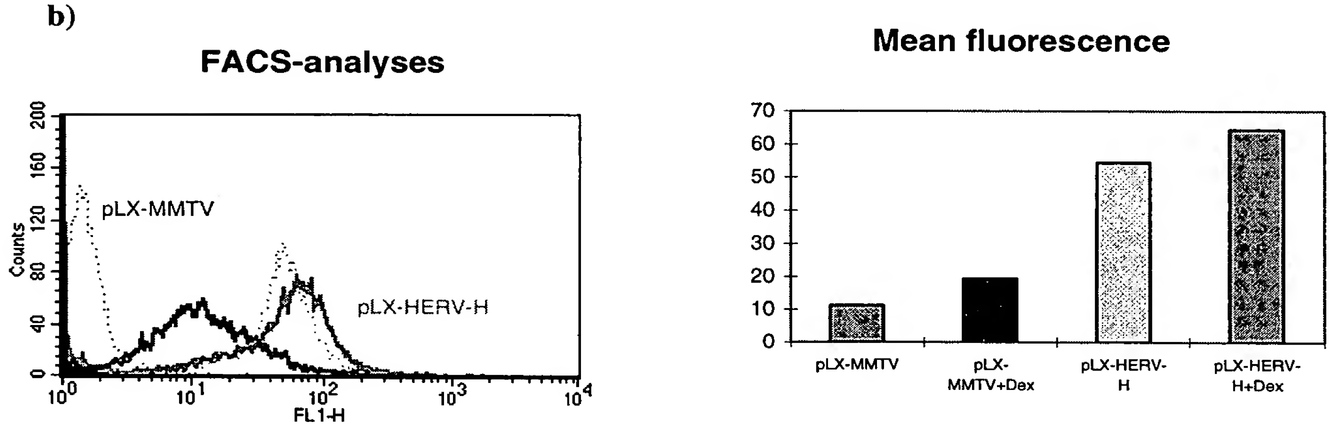



**Fig.10:** a) organization of the two ProCon vectors pLX-MMTV and pLX-HERV-H
b) promoter activity of the HERV-H LTR in comparison to the MMTV-LTR by infection of CrfK cells

## Appendix

### A. HERV-H LTR sequences

```
                  1                                                    50
HERV-H L31        TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCT GTGACCTGCA
HERV-H HCM2       TGTCAGGCCT CTGAGCCCAA GCCAGGCCAT CGCATCCCCT GTGACTTGCA
HERV-H 19         TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCT GCGACCTGCA
HERV-H MP20       TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCT GTGACCTGCA
HERV-H CM3        TGTCAGGCCT CTGAGCCCAA GCCAAGCCAT CGCATCCCCT GTGACTTGCA
HERV-H MC16       TGTCAGGCCT CTGAGCCCAA GCC....... .......... ......TGCA
HERV-H CM1        TGTCAGGCCT CTGAGCCCAA GCCAAGCCAT CGCATCCCCT GTGACTTGCA
HERV-H MP23       TGTCAGGCCT CTGAGCCCAA GCCAAGCCAT CGCATCCCCT GTGACTTGCA
HERV-H H13        TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCA GGGACCTGCA
HERV-H H1         TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CAAATCCCCT GTGACCTGCA
HERV-H HU8        TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCC.. GTGACCTGCA
HERV-H PA7        TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CAAATCCCCT GTGACCTACA

                  51                                                  100
HERV-H L31        CGTATACATC CAGATAGCC. .....TGAAG CAACTG.... ..........
HERV-H HCM2       CGTATACATC CAGATGGCC. .....TAAAG TAACTGAAGA ......TCCA
HERV-H 19         CATATACATC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H MP20       CGTACACATC CAGATGGCCG GTTCCTGCCT TAACTGATGA CATTCCACCA
HERV-H CM3        CGTGTATGCC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H MC16       CGTATACATC CAGATG.... .......AAG CAAGTGAAGA ......ATCA
HERV-H CM1        CGTATACGCC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H MP23       CGTATACGCC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H H13        CGTATACATC CAGATGGCC. .....TGAAG CAACTGAAGA ......TCCA
HERV-H H1         GGTGTACATC CAGATGACC. .....TGAAG CAACTGAAGA ......TCCA
HERV-H HU8        ..TATACATC CAGATGGCC. .....TGAAG CAACTGAAGA ......TCCA
HERV-H PA7        CGTGTACATC CAGATGACC. .....TGAAG CAACTGAAGA ......TCCA

                  101                                                 150
HERV-H L31        .........T AAAAATATCC TTAACTGATG ACA....... ..TTCCAATA
HERV-H HCM2       CAAAAGAAGT AAAAACAGCC TTAACTGATG ACA....... ..TTCCAACA
HERV-H 19         CAAAAGAAGT GAAAATGGCC TGTTCC.... .......... ..........
HERV-H MP20       CGAAAGAAGT GAAAATGACC TGTTCG.... .......... ..........
HERV-H CM3        CAAAAGAAGT GAAAAGGCCC TGCCCC.... .......... ..........
HERV-H MC16       CAAAAGAAGT GAAAATGGCC GGTTCC.... .......... ..........
HERV-H CM1        CAAAAGAACT GAAAAGGCCC TGCCCCGCCT TAACTGATGA CATTCCACCA
HERV-H MP23       CAAAAGAACT GAAAAGGCCC TGCCCCGCCT TAACTGATGA CATTCCACCA
HERV-H H13        CAAAGGAAGT GAAAATAGCC TTAACTGATG ACA....... ..TTCCACCA
HERV-H H1         CAAAAGAAGT GAAAGTAGCC TTAACTGATG ACA....... ..TTCCACCA
HERV-H HU8        CAAAGGAAGT GAAAATAGCC TTAACTGATG ACA....... ..TTCCACCA
HERV-H PA7        CAAAAGAAGT GAAAGTAGCC TTAACTGATG ACA....... ..TTCCACCA

                  151                                                 200
HERV-H L31        TTGTGATTTG TTTCTGCCCT ACCCTGACTG ATCAATGTGC TTTGTAATCT
HERV-H HCM2       TTGTGATTTG TTCCTGCCCC ACCCTAACTG ATAAATGTAC TTTGTAATCT
HERV-H 19         .......... .........T GCCTTAACTG ATGACATTAC CTTGTGAAAT
HERV-H MP20       .......... .........T GCCTTAACTG ATGACATTGT CTTGTGAAAT
HERV-H CM3        .......... .......... ACCTTAACTG AGTGATTAAC CCCATGAATT
HERV-H MC16       .......... .........T GCCTTAACTG ATGACATTAC CTTGTGAAAT
HERV-H CM1        TGGTGATTTG TTCTTGCCCC ACCTTAACTG AGTGATTAAC CCTGTGAATT
HERV-H MP23       TGGTGATTTG TTCTTGCCCC ACCTTAACTG AGTGATTAAC CCTGTGAATT
HERV-H H13        TTGTGATTTG TTTCTGCCCC ATCCTAACTG ATCAATGTAC TTTGTAATCT
HERV-H H1         TTGTGATTTG TTCCTGCCCC ACGCTAACTG AT......AC CATATATTCT
HERV-H HU8        TTGTGATTTG TTCCTGCCCC ACGCTAACTG AT......AC CATATATTCT
HERV-H PA7        TTGTGATTTG TTCCTGCCCT ACGCTAGCTG AT......AC CATATATTCT

                  201                                                 250
HERV-H L31        CCCCCACCCT TCAGAAGGCT CTTTGTAATC CTCCCCACCC TTGAGAATGG
HERV-H HCM2       CCCCCACCCT TAAGAAGGTC CTTTGTAATT CTCCCCACCC TTGAGAGTGT
HERV-H 19         TCCTTCTCCT GGCTCATCCT GGCTCAAAAG CTC..CCGCA CTGAG....C
HERV-H MP20       TCCTCCTCCT GGCTCATCCT GGCTCAAAAG CTC..CCCGA CTGAG....T
HERV-H CM3        TCCTTCCCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
HERV-H MC16       TCCTTCTCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
```

```
HERV-H CM1     TGCTTCTCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
HERV-H MP23    TGCTTCTCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
HERV-H H13     CTCCCACCCT TAAGAAGGTT CTTTGTAATT CTCCCCACCC TTGAGAGTGT
HERV-H H1      TCCCC..... .......... .......... .....CGCCC TTGAGAATGT
HERV-H HU8     TCCCC..... .......... .......... .....CGCCC TTGAGAATGT
HERV-H PA7     TCCCC..... .......... .......... .....CGCCC TTGAGAATGT

               251                                            300
HERV-H L31     ACTTGGTGAG ATCCACCCCC TGCCTGCAAA GCATTGCCCC TAACTCCACC
HERV-H HCM2    ACTTTGTGAG ATCCACACCT GCCCACCAGA GAACAAACCC CCTTTGACTG
HERV-H 19      ACCTTGTGAC CCCTGCCTCT GCCCGCCAGA GAGCAACCCC CTCTTGACTG
HERV-H MP20    ACATTGTGAC CCCCACTCCT GCCCGCCAGA GAACAGCCCC CT.TTGACTG
HERV-H CM3     ACCTTGTGAC CCCTGCCCCT GCCCACCAGA GAACAACCCC CT.TTGACTG
HERV-H MC16    ACCTTGTGAC CCCCACTCCT CCCCGCCACA GAACAACCCC CT.TTGACTG
HERV-H CM1     ACCTTGTGAC CCCCGCCCCT GCCCACCAGA GAACAACCCC CT.TTGACTG
HERV-H MP23    ACCTTGTGAC CCCCGCCCCT GCCCACCAGA GAACAGACCC CT.TTGACTG
HERV-H H13     ACTTTGTGAG ATCCACCCCC TGCCGGCAAA ACATTGCTCC TAACCCAACC
HERV-H H1      ACTTTGTA.. .......... .......... .......... .........C
HERV-H HU8     ACTTTGTA.. .......... .......... .......... .........C
HERV-H PA7     ACTTTGTA.. .......... .......... .......... .........C

               301                                            350
HERV-H L31     GCCTGTCCCA AAACCTATGA GAA.CTAATG ATA....... ATCCC.ACCA
HERV-H HCM2    TAATTTTCCA TTACCTTCCC TAATCCTATA AAACGGCCCC ACCCC.ATCT
HERV-H 19      TAATTTTCCT TTACCTACCT AAATCCTATA AAATGGCCCC ACTCCTATCT
HERV-H MP20    TAATTTTCCT TTATCTACCC AAATCCTATA AAACAGCCCC ACCTTTATCT
HERV-H CM3     TAATTTTCCA TTACTTTCCC AAATCCTATA AAACGGCCCC ACCCCTATCT
HERV-H MC16    TAATTTTCCA CTGCCCGCCC AAACCCTATA AAACGGTCCC ACCCC.ATCT
HERV-H CM1     TAATTTTCCA TTACCTTCCC AAATCCTATA AAACGGCCCC ACCCCTATCT
HERV-H MP23    TAATTTTCCA TTACCTTCCC AAATCCTATA AAACGGCCCC ACCCCTATCT
HERV-H H13     GCCTA.CCCC AAACCTGTAA GAA.CTAATG ATA........ATCC..ACCA
HERV-H H1      ACCTATCCC. AAACCTATAA GAA.CTAATG ATA....... ATCCT.ACCA
HERV-H HU8     ACCTATCCC. AAACCTATAA GAA.CTAATG ATA....... ATCC..ACCA
HERV-H PA7     ACCTATCCC. AAACCTATAA GAA.CTAATG ATA....... ATCCT.ACCA

               351                                            400
HERV-H L31     CACTTTGCTG ACTCTCTTTC C...AGACTC AGCCCGGCTG CACCCAGGTG
HERV-H HCM2    CCCTTTGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H 19      CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG TACCCAGGTG
HERV-H MP20    CCCTTGGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H CM3     CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H MC16    CCCTTCCCTG ACTCTCTTTT CTTCGGACTC AGCCCGCCTG CACCCAGGTG
HERV-H CM1     CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CCCCCAGGTG
HERV-H MP23    CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CCCCCAGGTG
HERV-H H13     CCCTTTGCTG ACTC..TTTT C...AGAATC AGCCCGCCTG CACCCAGGTG
HERV-H H1      CCCTTTGCTG ACTCTCTTTT T...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H HU8     CCCTTTGCTG ACTCTCTTTT T...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H PA7     CCCTTTGCTG ACTCTCTTTT T...GGACTC AGCCCGCCTG CACCCAGGTG

               401                    425
HERV-H L31     AAATAAACAG CCATGTTGCT CACAT
HERV-H HCM2    AAATAAACAG CCATGTTGCT CACAT
HERV-H 19      AAATAAACAG CCATGTTGCT CACAT
HERV-H MP20    AAATAAACAG CCATGTTGCT CACAT
HERV-H CM3     AAATAAACAG CCATGTTGCT CACAT
HERV-H MC16    AAATAAACAG CCATGTTGCT CACAT
HERV-H CM1     AAATAAACAG CCATGTTGCT CACAT
HERV-H MP23    AAATAAACAG CCATGTTGCT CACAT
HERV-H H13     AAATAAACAG CCATGTTGCT CACAT
HERV-H H1      AAATAAACAG CCATGTTGCT CACAT
HERV-H HU8     AAATAAACAG CCATGTTGCT CACAT
HERV-H PA7     AAATAAACAG CCATGTTGCT CACAT
```

**B. HERV-W LTR sequences**

```
                1                                                    50
HERV-T47D-W2    TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCGA
HERV-T47D-W4    TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCGA
HERV-T47D-W5    TGTTGAGATG GGGGACTGAG AGACAGGACT ACCTGGATTT CCTAGGCCGA
HERV-W1         TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCAA
HERV-W10        TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTC CCTAGGCCGA
HERV-W11        TGTTGAGATG GGGGACTGAG AGACAGGACT AGTTGGATTT CCTAGGCTGG
HERV-W18        TGTTGAGATG GGGGACTGAG AGACAGGACT AGTTGGATTT CCTAGGCCGG
HERV-W2         TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCAA
HERV-W22        TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCTGA
HERV-W23        TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCTGA
HERV-W4         TGTTGAGATG GGGGACTGAG AGACAGGACT AGTTGGATTT CCTAGGCTGG
HERV-W5         TGTTGAGATG GGGGACTGAG AGACAGGACT ACCTGGATTT CCTAGGCCAA
HERV-W6         TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCAA
HERV-W8         TGTTGAGATG GGGGACTGAG AAACAGGACT AGCAGGATTT CCTAGGCCGA

                51                                                  100
HERV-T47D-W2    CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCGCATCC ACCTTTAAAC
HERV-T47D-W4    CTAAGAATTC CTAAGCCTAG CTGGGAAGGT GACCGCATCC ATCTTTAAAC
HERV-T47D-W5    CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCACATCC ACCTTTAAAC
HERV-W1         CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAAAC
HERV-W10        CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCACATCC ACCTTTAAAC
HERV-W11        CTAAGAATCC CTAAGCCTAG CTGGGAAATT GACCACGTCC ACCTTTAAAC
HERV-W18        CTAAGAATCC CTAAGCCTAG CTGGGAAATT GACCACGTCC ACCTTTAAAC
HERV-W2         CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAACC
HERV-W22        CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCGCATCC ATCTTTAAAC
HERV-W23        CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAACC
HERV-W4         CTAAGAATCC CTAAGCCTAG CTGGGAAATT GACCACGTCC ACCTTTAAAC
HERV-W5         CTAAGAATCT CTAAGCCTAG CTGGGAAGGT GACCACATCC ACCTTTAAAC
HERV-W6         CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAACC
HERV-W8         TTAAGAATCC CTAAGCCTAG ATGGGAAGTT GACCACATCC ACCTTTAAAC

                101                                                 150
HERV-T47D-W2    ACGGGGCTCG CAACTTAGCT CACACCCAAC CAATCAGGTA GTAAAGAGGG
HERV-T47D-W4    ATGGGGCTTG CAACTTAACT CATATCTGAC CAATCAGGTA GTAAAGAGAG
HERV-T47D-W5    ACAGGGCTTG CAACTTAGCT CACACTTGAC CAGTCAGGTA GTAAAGAGAG
HERV-W1         ATGGGGCTTG CAACTTAGCT CACACCCAAC CAATCAGGTA GTAAAGAGAG
HERV-W10        ACGGGGCTTG CAACTTAGCT CATACCCAAC AAATCAGGTA GTAAAGAGAG
HERV-W11        ACGGGGCTTG CAATTTAGCT CACACCCGAC CAATCAGGTA GTAAAGAGAG
HERV-W18        ACGGGGCTTG CAATTTAGCT CACACCCGAC CAATCAGGTA GTAAAGGGAG
HERV-W2         ACTAGGCTTG CAACTTAGCT CACACCCGAC CAATCAGGTA GTAAAGGGAG
HERV-W22        ATGGGGCTTG CAACTTAACT CATATCTGAC CAATCAGGTA GTAAAGAGAG
HERV-W23        ACTAGGCTTG CAACTTAGCT CACACCCGAC CAATCAGGTA GTAAAGAGAG
HERV-W4         ACGGGGCTTG CAATTTAGCT CACACCCGAC CAATCAGGTA GTAAAGGGAG
HERV-W5         ACAGGGCTTG CAACTTAGCT CACACCCGAC CCATCAGGTA AGAAAGAGAG
HERV-W6         ACTAGGCTTG CAACTTAGCT CACACCCGAC CAATCAGGTA GTAAAGAGAG
HERV-W8         ACGGGGCTTG CAACTCAGCT CACACCCGAC CCATCAGGTA AGAAAGAGAG

                151                                                 200
HERV-T47D-W2    CTCACTAAAA TGCTAATTAG GCAAAGACAG GAGGTAAAGA AATAGCCAAT
HERV-T47D-W4    CTCACTAAAA TGCTAATTAG GCTAAAACAG GAGGCAAAGA AGTAGCCAAT
HERV-T47D-W5    CTCACTAAAA TGCTAATTAG GCTAAAACAG GAGGTAAAGA AATAGACAAT
HERV-W1         CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W10        CTCACTAAAA TACTGATTAG GCGAAAACAG GAGGTAAGGA AACAGCCAGT
HERV-W11        CTCACTAAAA TGCTAATTAG GGAAAAACAG GAGGTAAAGA AGTAGCCAAT
HERV-W18        CTCACTAAAA TGCTAATTAG GGAAAAACAG GAGGTAAAGA AGTAGCCAAT
HERV-W2         CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAGAGA AATAGCCAAT
HERV-W22        CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W23        CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W4         CTCACTAAAA TGCTAATTAG GGAAAAACAG GAGGTAAAGA AGTAGCCAAT
HERV-W5         CCCGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGTCAAT
HERV-W6         CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W8         CCCGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAAT
```

```
              201                                                  250
HERV-T47D-W2  CATCTATTGC CTGAGAGCAC AGCAGGAGGG ACAATGATCG GGATATAAAC
HERV-T47D-W4  CATCTGTTGC CTGACAGCAC AGCAGGAGGG ACAATGATCG GGATATAAAC
HERV-T47D-W5  CATCTATCAC CTGAGAGCAC AGTGGGAGGG ACAATGATCG GCATATAAAC
HERV-W1       CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W10      CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W11      CATCTATCGC CTGAGAGCAC AACAGGAGGG ACAATGATCA GGATATAAAC
HERV-W18      CATCTATCGC CTGAGAGCAC AACAGGAGGG ACAATGATCA GGATATAAAC
HERV-W2       CATCTATCGC CTGAGAGCAC AGCAGGAGGG ACAATGATCC GGATATAAAC
HERV-W22      CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W23      CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W4       CATCTATCGC CTGAGAGCAC AACAGGAGGG ACAATGATCA GGATATAAAC
HERV-W5       CATCTATTGC CTGAGAGCAC AGCGGGAGGG ACAATGATCA GGATATAAAC
HERV-W6       CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W8       CATCTATTGC CTGAGAGCAC AGCGGGAGGG ACAATGATCA GGATATAAAC

              251                                                  300
HERV-T47D-W2  CCAAGTCTTC GAGCCGGCAA TGGCTACCTT CTTTGGGTCC CCTCCCTTTG
HERV-T47D-W4  CCAGGCATTC GAGCCAGCTA CAGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-T47D-W5  CCAGGCATTC GAGCCAGCAA CAGCAACCCC CTTTGGG... ..........
HERV-W1       TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W10      TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W11      CCAGGCATTC AAGCCAGCGG TGGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W18      CCAGGCATTC AAGCCAGCGG TGGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W2       CCAAGCATTC GAGCCAGCAA TGGCTACCCT CTTTGTGTCC CCTCCCTTTG
HERV-W22      TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W23      TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W4       CCAGGCATTC AAGCCAGCGG TGGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W5       CCAGGCATTC GAGCCGGCAA CGACTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W6       TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W8       CCAGGCATTC GAGCCGGCAA CGACTACCCT CTTTGGGTCC CCTCCCTTTG

              301                                         343
HERV-T47D-W2  TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG C..
HERV-T47D-W4  TATGGGAGCT CTGTCTTCAC TCTATTAAAT CTTGCAACTG C..
HERV-T47D-W5  ......AGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG C..
HERV-W1       TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W10      TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W11      TATGGGAGCC CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W18      TATGGAAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W2       TATGGGAGCT CTATTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W22      TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W23      TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W4       TATGGAAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W5       TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W6       TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W8       TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
```

## C. HERV-K LTR sequences

| | 1 | | | | 50 |
|---|---|---|---|---|---|
| **HERV-K45** | GCGACCGGT: | GGATC:CCGG | GCCCGCGG:T | ACCGTCGACT | :GCAGAATTC |
| **HERV-K27** | GCGACCGGT: | GGATC:CCGG | GCCCGCGG:T | ACCGTCGACT | :GCAGAATTC |
| **HERV-K2** | GCGACCGGT: | GGATC:CCGG | GCCCGCGG:T | ACCGTCGACT | :GCAGAATTC |
| **HERV-K1** | GCGACCGGT: | GGATC:CCGG | GCCCGCGG:T | ACCGTCGACT | :GCAGAATTC |
| **HERV-K30** | GTC | CCACCTCCAG | CCCTAAGGCG | GTTTTTCCCT | ATCTCAGTAG |
| **HERV-K10** | | | | | AGTAG |

| | 51 | | | | 100 |
|---|---|---|---|---|---|
| **HERV-K45** | ATGGAGCATA | CAATCGGGTT | TTATACCGAG | ACATTCCATT | GCCCAGGGAC |
| **HERV-K27** | ATGGAGCATA | CAATCGGGTT | TTATACCGAG | ACATTCCATT | GCCCAGGGAC |
| **HERV-K2** | ATGGAGCATA | CAATCGGGTT | TTATACCGAG | ACATTCCATT | GCCCAGGGAC |
| **HERV-K1** | ATGGAGCATA | CAATCGGGTT | TTATACCGAG | ACATTCCATT | GCCCAGGGAC |
| **HERV-K30** | ATGGAGCATA | CAATCGGGTT | TTATACCGAG | ACATTCCATT | GCCCAGGGAC |
| **HERV-K10** | ATGGAGCATA | CAATCGGGTT | TTATACCGAG | ACATTCCATT | GCCCAGGGAC |

| | 101 | | | | 150 |
|---|---|---|---|---|---|
| **HERV-K45** | AGGCAGGAGA | CAGATGCCTT | CCTCTTGTCT | CAACTGCAAG | AGGCATTCCT |
| **HERV-K27** | AGGCAGGAGA | CAGATGCCTT | CCTCTTGTCT | CAACTGCAAG | AGGCATTCCT |
| **HERV-K2** | AGGCAGGAGA | CAGATGCCTT | CCTCTTGTCT | CAACTGCAAG | AGGCATTCCT |
| **HERV-K1** | AGGCAGGAGA | CAGATGCCTT | CCTCTTGTCT | CAACTGCAAG | AGGCATTCCT |
| **HERV-K30** | AGGCAGGAGA | CAGATGCCTT | CCTCTTGTCT | CAACTGCAAG | AGGCATTCCT |
| **HERV-K10** | AGGCAGGAGA | CAGATGCCTT | CCTCTTGTCT | CAACTGCAAG | AGGCATTCCT |

| | 151 | | | | 200 |
|---|---|---|---|---|---|
| **HERV-K45** | TCCTCTTATA | CTAATCCTCC | TCAGCACAGA | CCCTTTACGG | GTGTCGGGCT |
| **HERV-K27** | TCCTCTTATA | CTAATCCTCC | TCAGCACAGA | CCCTTTACGG | GTGTCGGGCT |
| **HERV-K2** | TCCTCTTATA | CTAATCCTCC | TCAGCACAGA | CCCTTTACGG | GTGTCGGGCT |
| **HERV-K1** | TCCTCTTATA | CTAATCCTCC | TCAGCACAGA | CCCTTTACGG | GTGTCGGGCT |
| **HERV-K30** | TCCTCTTATA | CTAATCCTCC | TCAGCACAGA | CCCTTTACGG | GTGTCGGGCT |
| **HERV-K10** | TCCTCTTTTA | CTAATCCTCC | TCAGCACAGA | CCCTTTACAG | GTGTCGGGCT |

| | 201 | | | | 250 |
|---|---|---|---|---|---|
| **HERV-K45** | GGGGGACGGT | CAGGTCTTTC | CCTTCCCACG | AGGCCATATT | TCAGACTATC |
| **HERV-K27** | GGGGGACGGT | CAGGTCTTTC | CCTTCCCACG | AGGCCATATT | TCAGACTATC |
| **HERV-K2** | GGGGGATGGT | CAGGTCTTTC | CCTTCCCACG | AGGCCATATT | TCAGACTATC |
| **HERV-K1** | GGGGGACGGT | CAGGTCTTTC | CCTTCCCACG | AGGCCATATT | TCAGACTATC |
| **HERV-K30** | GGGGGACGGT | CAGGTCTTTC | CCTTCCCACG | AGGCCATATT | TCAGACTATC |
| **HERV-K10** | GGGGGACGGT | CAGGTCTTTC | CCTTCCCACG | AGGCCATATT | TCAGACTATC |

| | 251 | | | | 300 |
|---|---|---|---|---|---|
| **HERV-K45** | ACATGGGGAG | AAACCTTGGA | CAATACCTGG | CTTTCCTAGG | CAGAGGTCCC |
| **HERV-K27** | ACATGGGGAG | AAACCTTGGA | CAATACCTGG | CTTTCCTAGG | CAGAGGTCCC |
| **HERV-K2** | ACATGGGAAG | AAACCTTGGA | CAATACCTGG | CTTTCCTAGG | CAGAGGTCCC |
| **HERV-K1** | ACATGGGGAG | AAACCTTGGA | CAATACCTGG | CTTTCCTAGG | CAGAGGTCCC |
| **HERV-K30** | ACATGGGGAG | AAACCTTGGA | CAATACCTGG | CTTTCCTAGG | CAGAGGTCCC |
| **HERV-K10** | ACATGGGGAG | AAACCTTGGA | CAATACCTGG | CTTTCCTAGG | CAGAGGTCCC |

| | 301 | | | | 350 |
|---|---|---|---|---|---|
| **HERV-K45** | TGCGGCCTTC | CGCAGTTTTT | GTGT:CCTGG | GTACTTGAGA | TTAGGGAGTG |
| **HERV-K27** | TGCGGCCTTC | CGCAGTTTTT | GTGT:CCTGG | GTACTTGAGA | TTAGGGAGTG |
| **HERV-K2** | TGCGGCCTTC | CGCAGTTTTT | GTGT:CCTGG | GTACTTGAGA | TTAGGGAGTG |
| **HERV-K1** | TGCGGCCTTC | CGCAGTTTTT | GTGT:CCTGG | GTACTTGAGA | TTAGGGAGTG |
| **HERV-K30** | TGCGGCCTTC | CGCAGTTTTT | GTGTCC:TGG | GTACTTGAGA | TTAGGGAGTG |
| **HERV-K10** | TGCGGCCTTC | TGCAGTTTTT | GTGTCCCTGG | GTACTTGAGA | TTAGGGAGTG |

| | 351 | | | | 400 |
|---|---|---|---|---|---|
| **HERV-K45** | GTGATGACTC | TTAAGGAGCA | TGCTGCCTTC | AAGCATCTGT | TTAACAAAGC |
| **HERV-K27** | GTGATGACTC | TTAAGGAGCA | TGCTGCCTTC | AAGCATCTGT | TTAACAAAGC |
| **HERV-K2** | GTGATGACTC | TTAAGGAGCA | TGCTGCCTTC | AAGCATCTGT | TTAACAAAGC |
| **HERV-K1** | GTGATGACTC | TTAAGGAGCA | TGCTGCCTTC | AAGCATCTGT | TTAACAAAGC |
| **HERV-K30** | GTGATGACTC | TTAAGGAGCA | TGCTGCCTTC | AAGCATCTGT | TTAACAAGGC |
| **HERV-K10** | GTGATGACTC | TTAAGGAGCA | TGCTGCCTTC | AAGCATCTGT | TTAACAAAGC |

```
           401.............................................450
HERV-K45   ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K27   ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K2    ACATCCTGCA CTGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCGCAC
HERV-K1    ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K30   ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K10   ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAT

           451.............................................550
HERV-K45   GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K27   GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K2    GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K1    GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K30   GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K10   GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC

           501.............................................550
HERV-K45   AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K27   AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K2    AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K1    AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K30   AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K10   AAGGCAGAAG AATTTTTCTT AGCACATAAC AAAATGGAGT CTCCTATGTC

           551.............................................600
HERV-K45   TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K27   TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K2    TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K1    TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K30   TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCCCTC TTGCTTTTCC
HERV-K10   TACTTCTTTC TACACAGACA CAGTAACAAT TTGATCTCTC TTGCTTTTCC

           601.............................................650
HERV-K45   CCACATTTCC CCCTTTTCTT TTCG
HERV-K27   CCACATTTCC CCCTTTTCTT TTCGA
HERV-K2    CCACATTTCC CCCTTTTCTT TTCGACAAA
HERV-K1    CCACATTTCC CCCTTTTCTT TTCGACAAAA CCGCCAT:CT CGAGATC:TG
HERV-K30   CCACATTTCC CCCTTTTCTT ATCCATCACA CTGGCGGCCG CTCGAGCATG
HERV-K10   CCACATTTCC CCCTTTTCTT TTCGACAAAA CCGCCATC

           651...........................
HERV-K1    AGT
HERV-K30   CATCTAGAGG GCCCAATTCG CCCTATAGTG
```

**HERV-K-T47D-5'LTR**

TGTGGGCGAAGGATTACCCAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATATAGAAAATAGCTAG
AATAAGAATAGTTATAATAAAAATTAGATATACACATGATCATGGACATTACCAATCATTACTACAAACATTGTTAATCA
TTAGCTTTTAATATTACTCTTTGTTTTATTACTAATATAACCAAGGAATAACCGGTAGCATACGGTCAGGTGCTGAAGGG
ACATTGTGAGAAGTGACCTAGAAGGCAAGAGGTGAGCCTTCTGTCACGCCTGCATAAGGACAGCTTGAGGGCTCCTTGGT
CAAGCTGTAACACCAGTGCCTGGGAAGGCACCGTTACTTAGCAGACCATGAAAGGGAGTCTCCATTCCTTGGAGGAGTCA
GGGAAACACTATGCTCCACCAGCTTCTTGTGTATCCAGCCCTGCCCACAGTCATCCAGAGGCATAAACCCCTCCCTGTGG
TGCTGTGCTTCAATGGCCATGCTTCTTGTCCACTTTCATGTTCCTCCTGTACTCCTGGTTCCTCTTTGAAGTTCGTAGAA
GATAATGGTAGAAGAAATAGTGAAAGTCTTTGATCTTTCTTATAAGTGCATAGAAGAAAACACTGATGTATGCCTGCCTT
CCCTCTCTGCTTCAGCTACCTAAAAGGAAAGGCCCCCTTTCCCATGATCACATGACTTGCCTGACCTTATCAATCACTTG
GAGGACTCACCCTCCTTACCCTGTCCCTTTGTCTTGTATGCAATAAATATCAGCACGCCCAGCCATTCGGGGCCACTACT
GGTCTCCGCAACTTGGTGGTAGTGGTACCCTGGGCCCAGCTGTTTTCTCTTTATCTCTTTTGTCTTGTGTCTTTATTTCT
TACAATCTCTCATCTCTGCACATGGGGAGAACACCGGCAAAGCCCGTAGGGCTGGACCTTACA

**L48-LTR (U3-R)**

TGTGGGCGGAAGAGTACCTAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATAGAATA
AGAATAGTCATAATACAAATTAGATACAGCGATGATCATGAACAATTATCCATCATTATTATAAACATTATTAATCATTA
GCTTTTAATATTACTCTGTTGCATTAATAATATAACCTAGGAATAACCGGCAGGTATAGGGTCAGGTGCTGAAGGGACAT
TGTGAGAAGTGAATAGAAGGCAAGAGGGGAGCCTTCTGTCATGCCCGCATAAGGGCCGCTTGAGGGCCCCTTGGTCAAGC
GGTAACGCCAGTGTCTGGGAAGGCACCCGTTACTGAGCAGACCGGGAAAGGGAGTCTCCTTTCCTTGGAGGAGTCAGGGA
ACGCTCTGCTCCACCAGCTTCTTGTGGGAGGCTGGATGTTACCCAGGCCTGCCTGCAGTCATCCGGAGGCCTGAACCCCT
CCCTGTGGTGCTTCAATGGTCACGTTCCTTGTCCACTTTCATGCTCCTTCCGTACTCCTGGTTCCTCTTTGAAGTTCGTA
GTAGATAGCGGTAGAAGAAATAGTGAAAGTCTTAAAGTCTTTGATCTTATAAGTTCATAGAAGAAAACGCTGATGCCTGC
CGCCTTCTCTCTCTGCTTCAGCTACCTAAGAGGGAAGGGCCCGCTGTCCTGTGATCAGGTGACTTGCTTCACCTTGTCAA
TCACTTAGAAGACTGACCCTCCTTATCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGCGAGCCCAGCCGTTCAGGGC
CACTACCGGTCTCCGTGTCTTTGTGGTAGTGGTCCCCGGGCCCAGCTGTTTTCTCTTT

**L5-LTR (U3-R)**

TGTGGGTGGAGGATTACCCAGGTGCCAAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAAAAAATAGAATA
AGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTAGCAATCACTATTAATCTTTAGCTTTTAATATT
ACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAGGGACATTGTGTGAAGT
GACCTGGAAGGCAAGAGGTGAGCCCTCTGTCACGCCCACATAAGGGCCGCTTGAGGGCTCCTTGGTCAAGTGGTAACGCC
AGTGTCTGGGAATGCACCCGTTAATTAGCAGACCGCGAAAGGGAGTCTCCTTTCCTTGGAAGAGTTGGGGAACACTCTGC
TCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCGCGCAGTCATCCGGAGGCTTAAACCCTTCCCTGTGGT
GCTGTGCTTCAATGGTCCCACTCCTTGTCCACTTTCATGCTCCTCCCGTACTCCTGGTTCCTCTTTGAAGAGCGCAGTAG
ATAGCGGTAGAAGAAATAGTGAAAGTCTTAAAGTCTTCGATCTTTCTTACAAGTGCAGAGAAGAAAACGCTGACATATGC
TGCCTTCCCTCTCTGCTTCGGCTACCTAAAAGGGAAGGGCCGCCTATCCTGTAATCACATGACTTGCTTCACCTTGTCAA
TCACTTAGAAGATTCACTCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGACCCCAGCCGTTCAGGGC
CACTACTGGTCTCCGCGTCTTGATGGTAGTGGTCACCCCGGCCC

**L50-LTR (U3-R)**

TGTGGGTGGAGGATTACCCAGGTGCCAAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAAAAAATAGAATA
AGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTAGCAATCACTATTAATCTTTAGCTTTTAATATT
ACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAGGGACATTGTGAGAAGT
GACCTGGAAGGCAAGAGGTGAGCCCTCTGTCACGCCCACATAAGGGCCGCTTGAGGGCTCCTTGGTCAAGTGGTAACGCC
AGTGTCTGGGAATGCACCCGTTAATTAGCAGACCGCGAAAGGGAGTCTCCTTTCCTTGGAAGAGTTGGGGAACACTCTGC
TCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCGCGCAGTCATCCGGAGGCTTAAACCCTTCCCTGTGGT
GCTGTGCTTCAATGGTCCCACTCCTTGTCCACTTTCATGCTCCTCCCGTACTCCTGGTTCCTCTTTGAAGAGCGCAGTAG
ATAGCGGTAGAAGAAATAGTGAAAGTCTTAAAGTCTTCGATCTTTCTTACAAGTGCAGAGAAGAAAACGCTGACATATGC
TGCCTTCCCTCTCTGCTTCGGCTACCTAAAAGGGAAGGGCCGCCTATCCTGTAATCACATGACTTGCTTCACCTTGTCAA
TCACTTAGAAGATTCACCCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGACCCCAGCCGTTCAGGGC
CACTACTGGTCTCCGCGTCTTGATGGTAGTGGTCACCCCGGCCCAGGTGTTTTTTCTTT

**L9-LTR (966 nt)**

TGTGGGTGGAGGATTACCCAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATGGTTAG
AATAAGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTATCAATCATTATTATAAACATTATTAATC
ATTAGCTTTTAATATTACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAA
GGACATTGGGAGAAGTGACCTAGAAGGCAAGAGGTGAGTCTTCTGTCACGCCCGCATAAGGGTTGCTTGAGGGCTCCTTG
GTCAAGTGGTAACGCCGGTGTCTGGGAAGGCACCTGTTACTTAGCCGACCACGAAAGGGAGTCTCCTTTCCTTGGAGGAG
TCAGGGCGCACTCTGCTCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCCCGCAGTCATCCGGAGGCCTA
AACCCCTCCCTGTGGTGCTGTGCTTCAATGGGCACACTCCTCGTCCACTTTCATGTTCCTCCCATACTCCTGGTTTCTCT
TTGAAGTTCGTAGTAGATAGTGGTAGAAGGAATAGGGAAAATCTTAAAGTGTTTGATCTTTCTTATAAGTGCATAGAAGA
AAACGCTGACATATGCTGCCTTCTCTGTCTGCTTCAGCTACCTAAGAGGGAAGGGCCCCCTGTCCAGTGATCACGTGACT
TGCTTCACCTTGTCAATCACTTAGAAGATTCACCCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGCA
CCCAGCCTTTCGGGGCCACTTACCGGTCTCCACGTCTTGGTGGTAGTGGTCCCCCGGGCCCAGCTGTTTTCTCTTTATCT
CTTTGTCTTGTGTCTTATTTATTACAATCTCTCGTCTCCGCACACAGGGAGAACACCCGCTAAGCTCCGTAGGGCTGGAC
CCTACA

**L8-LTR (938 nt)**

TGTGGGTGGAGGATTACCCAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATGGTTAG
AATAAGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTATCAATCATTATTATAAACATTATTAATC
ATTAGCTTTTAATATTACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAG
GGACATTGGGAGAAGTGACCTAGAAGGCAAGAGGTGAGTCTTCTGTCACGCCCGCATAAGGGTTGCTTGAGGGCTCCTTG
GTCAAGTGGTAACGCCGGTGTCTGGGAAGGCACCTGTTACTTAGCCGACCACGAAAGGGAGTCTCCTTTCCTTGGAGGAG
TCAGGGCACACTCTGCTCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCCCGCAGTCATCCGGAGGCCTA
AACCCCTCCCTGTGGTGCTGTGCTTCAATGGGCACACTCCTCGTCCACTTTCATGTTCCTCCCATACTCCTGGTTCCTCT
TTGAAGTTCGTAGTAGATAGTGGTAGAAGGAATAGGGAAAATCTTAAAGTGTTTGATCTTTCTTATAAGTGCATAGAAGA
AAACGCTGACATATGCTGCCTTCTCTGTCTGCTTCAGCTACCTAAGAGGGAAGGGCCCCCTGTCCAGTGATCACGTGACT
TGCTTCACCTTGTCAATCACTTAGAAGATTCACCCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGCA
CCCAGCCTTTCGGKKCACTTACCGGTCTCCACGTCTTGGTGGTAGTGGTCCCCCGGCCCAGCTGTTTTCTCTTTATCTCT
TTGTCTTGTGTCTTATTTATTACAATCTCTCGTCTCCGCACACAGGGAGAACACCCGC (Abbruch 26 nt vor Ende der LTR)

**L49-LTR = L20-LTR (963 nt)**

TGTGGGCGAAAGATTACCTAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATAGTTAA
AATAAGAATAGTTATAATACAAATTAGATATAGAGATGATCATGGACAATTATCAATCATTATTATAAACATTAATCATT
AGCTTTTAATATTACTCTTTGTTGCTTTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGTTGACGGGA
TATTGTGAGAAGTGACCTAGAAGGCAAGAGGTGAGCCTTCTGTCACGCCCACATAAGGGCCGCTTGAGGGCTCTTTGGTC
AAGTGGTAACGCCAGTGTCTGTGAAGGCACCTGTTACTTAGCAGACCGCGAAAGGGAGTCTCCTTTCCTTGGAGGAGTCA
GGGAACACTCTGCTCCACCAGCTTCTTGTGGAAGGCTGGATATTATCTAGGCCTGCCCGCAGTCATCTGGAGGCCTAAAC
CCCTCCCTGTGGTGCTGTGCTTCAGTGGTCACTCTCCTTGTCCACTTTCATGTTCCTCCCGTACTCCTGGTTCCTCTTTG
AAGTTCGTAGTAGATAGCAGTAGAAGAAATAGTGAAAGTCTTAAAGTATTTGATCTTTCTTATAAGTGCATAGAAGAAAA
CGCTGACATATGCTGCCTTCTCTATCTCTGCGGTGGCTACCTAAAAGGGAAGGGCCCCCTGTCCCATGATCATGTGACTT
GCTTCACCTTATCACTTAGAAGATTCATCCTCCTTACCCTGCGCCCCCTCGTCTTGTATGCAATAAATATCAGCACGCCC
AGTCGTTTGAGGCCACTGCCGGTCTCCGCGTCTTGGTGGTAGTGGTCCCCCGGGCCCAGCTATTGTCTCTTTATCTCTTT
GTCTTGTGTCTTTATTTATTACAATCTCTTGTCTCTGCACACAGGGAGAACACCTGCTAAGCCCCGTAGGACTGGACCCT
ACA

**HERV-IP-T47D**

TGTTCAATTCTTTGCCTTCTACTTTTAAACTTAACTTCCTCATAAAGCAACCTTTTTCAATCACCTGCTCCACTCTGACT
CATTCTGATCACCTGCTCCACCCTGACTCATTCCGATCACCTGATCCACTGTGACTCATTCCGATTACCCGCTCCACCCT
GACTCATTCTGATTCTGATTTCCTGCTCTGCCATAACCATTTTTCCCGCCAAACCACTCACCCTGTCACTCTCTTTAAAT
TAGCCAATTGGAATTAGTTTAGCCTGTGCGGTCTAACCCTAGCCAATAGGGGACTGACACAGCAGCAGGGGCCACATGTG
TCAGGAATAAGACCCCCTTCCCCTCCCTGTCCAGATGTGTGCTCACCATTGCTCCATCTGTGAGGGCACACCCTTCTATA
GAAGTAAATTGCCTTGCTGAGAAGAAAAAAAAGAACATTTTATATTCAAGTCCTATTTCTTTTGCTGCACCGAAACTTTA
TTTATAACA